

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : <b>C07K 3/00, 13/00, C07H 21/00 C12P 21/06, 21/02, 21/04 C12N 15/00</b>		A1	(11) International Publication Number: <b>WO 92/14748</b> (43) International Publication Date: <b>3 September 1992 (03.09.92)</b>
(21) International Application Number: <b>PCT/US92/01300</b>		(72) Inventors; and	
(22) International Filing Date: <b>20 February 1992 (20.02.92)</b>		(75) Inventors/Applicants (for US only) : <b>TERMAN, Bruce, Israel [US/US]; 17 Prospect Street, Monroe, NY 10950 (US). CARRION, Miguel, Eduardo [EC/US]; 26 Summit Avenue, Spring Valley, NY 10977 (US).</b>	
(30) Priority data: <b>657,236 22 February 1991 (22.02.91) US</b>		(74) Agent: <b>GORDON, Alan, M.; American Cyanamid Company, 1937 West Main Street, P.O. Box 60, Stamford, CT 06904 (US).</b>	
(60) Parent Application or Grant (63) Related by Continuation US Filed on <b>657,236 (CIP) 22 February 1991 (22.02.91)</b>		(81) Designated States: <b>AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, KR, LU (European patent), MC (European patent), NL (European patent), SE (European patent), US.</b>	
(71) Applicant (for all designated States except US): <b>AMERICAN CYANAMID COMPANY [US/US]; 1937 West Main Street, P.O. Box 60, Stamford, CT 06904 (US).</b>		Published <i>With international search report.</i>	

(54) Title: IDENTIFICATION OF A NOVEL HUMAN RECEPTOR TYROSINE KINASE GENE

(57) Abstract

A DNA sequence encoding a novel human growth factor receptor referred to as a type III receptor tyrosine kinase is described. The amino acid sequence of the receptor is also described. The receptor has a sequence which is similar to that of the kinase domains of known type III receptor tyrosine kinases, but which is unique in its kinase insert domain sequence. The receptor binds specifically to the vascular endothelial cell growth factor.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	MI	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark	MG	Madagascar		
ES	Spain				

5

IDENTIFICATION OF A NOVEL HUMAN  
RECEPTOR TYROSINE KINASE GENE

10

FIELD OF THE INVENTION

This invention relates to the DNA sequence  
encoding a novel human growth factor receptor which is  
15 a type III receptor tyrosine kinase. The receptor is  
referred to as Kinase insert Domain containing Receptor  
(KDR) and binds specifically to the growth factor  
vascular endothelial cell growth factor (VEGF). This  
invention also relates to the amino acid sequence of  
20 the receptor.

BACKGROUND OF THE INVENTION

Growth factors are small molecules which  
25 regulate normal cell growth and development through  
interaction with cell surface receptors. The receptors  
for a number of growth factors are referred to as  
tyrosine kinases; that is, binding of growth factor to  
the receptor stimulates an increased phosphorylation of  
30 tyrosine amino acids within the receptor; this in turn  
leads to cellular activation (Bibliography 1).

There is increasing evidence that genetic  
alterations affecting the expression of receptor  
tyrosine kinases (RTK) can contribute to the altered  
35 cell growth associated with cancer. This conclusion is

- 2 -

supported by the frequent identification of RTK as products of the oncogenes for many of the acutely transforming retroviruses (e.g., 2,3,4) and the overexpression of RTK in certain cancers (5). The identification of a novel RTK may lead to a better understanding of cell growth under both normal and transforming circumstances.

The amino acid sequence in the catalytic domain of all tyrosine kinases has been conserved (6). Detailed analysis of the amino acid sequences within the catalytic and noncatalytic domains of RTK indicates the existence of distinct structural subtypes. One group of RTK (designated type III) includes the ckit proto-oncogene and the receptors for platelet derived growth factor (PDGF) and colony stimulating factor-1 (CSF-1).

20 The most unusual feature of this subtype is that its catalytic (kinase) domain is interrupted by a long insertion sequence of 12-102 amino acids (the kinase insert domain). The two peptides constituting the kinase domain are conserved between the receptors, while the sequence of the kinase insert domain is unique for each receptor.

Several approaches have been tried in order  
to identify novel RTK, including low-stringency  
screening of cDNA libraries with previously  
characterized DNA probes (7). More recently, a  
technique has been developed that is capable of greatly  
facilitating the identification of novel genes for  
which some sequence data are known. The polymerase  
chain reaction (PCR) has been used to identify novel  
members of several gene families including those of  
guanine nucleotide regulatory proteins (8) and protein  
phosphatases (9). PCR has been used to identify novel  
tyrosine kinase genes (10), though the primers used in

- 3 -

that study were designed from DNA segments contained in all tyrosine kinases, rather than being specifically directed against RTK. It is a continuing goal to identify receptors for growth factors.

5           The elucidation of the growth factors, as well as their receptors, involved in regulating endothelial cell function is critical for the understanding of how new blood vessels are formed (angiogenesis). Angiogenesis plays a significant role  
10          in both normal and pathological events such as embryogenesis, progression of ocular diseases, and wound healing (11). In particular, angiogenesis is an important process for the growth of tumors (11).  
15          Angiogenesis is a complex process involving endothelial cell proliferation, migration, and tissue infiltration. These events are stimulated by growth factors which either (i) act directly on endothelial cells (12,13), or (ii) act indirectly by inducing host cells to release specific endothelial cell growth factors (11).  
20          One member of the first group is vascular endothelial cell growth factor (VEGF), also known as vascular permeability factor (14-16). Besides its angiogenic activity, VEGF displays the physiological function of increasing the permeability of capillary vessels to different macromolecules (14).  
25

#### SUMMARY OF THE INVENTION

30          The present invention relates to novel DNA segments which together comprise a gene which encodes type III RTK. The type III RTK encoded by the gene is designated the KDR protein (which stands for Kinase insert Domain containing Receptor). The KDR protein binds specifically to the growth factor VEGF (vascular endothelial cell growth factor).  
35

- 4 -

The DNA segments are identified and isolated through the use of PCR technology. The overall strategy is summarized as follows:

PCR is used to amplify the DNA segments corresponding to the kinase insert domains of type III receptor tyrosine kinase genes in an endothelial cell library designated HL10246 (Clontech Laboratories, Inc., Palo Alto, CA). Degenerate oligonucleotide primers are designed which are complementary to conserved tyrosine kinase domains flanking the kinase insert domains of known type III receptor tyrosine kinases. These primers are used in the PCR procedure. DNA probes, designed from the DNA sequence of the PCR product, are then used to identify cDNA clones of the receptor gene from the original cDNA library.

In particular, the present invention relates to specific oligonucleotides which, when used as primers for PCR, allow for the amplification of DNA segments corresponding to the kinase insert domains of type III RTK genes.

In a principal embodiment, the present invention is directed to three overlapping DNA segments (designated BTIII081.8, BTIII129.5 and BTIV169) which comprise the entire coding region of this novel gene, namely, 4,068 nucleotides extending to the 3' end.

These DNA segments are isolated from a human endothelial cell cDNA library and together comprise the gene coding for a novel type III receptor tyrosine kinase. The human gene containing these DNA segments is referred to hereinafter as KDR (which stands for Kinase insert Domain containing Receptor) or, alternatively, as kdp (which stands for Kinase insert Domain containing Protein). The use of the term KDR is intended to include any DNA segments which form the

- 5 -

human gene which encodes the novel type III RTK of this application.

The DNA segments embodied in this invention are isolated from human sources. The present invention comprises DNA segments, and methods for using these DNA segments, which allow for the identification of a closely related gene in mouse DNA. The methods developed in this invention can be readily used by those skilled in the art for the identification and isolation of closely-related homologues in other species. Therefore, the present invention also embodies all DNA segments from species other than human which encode proteins having substantially the same amino acid sequence as that encoded by the kdp gene.

The present invention further relates to methods developed for the detection of mRNA's produced as a result of transcription of the sense strands of the DNA segments of this invention. Messenger RNA prepared from bovine endothelial cells are used in developing these methods. The ability to detect mRNA for a novel RTK may ultimately have medical benefit, especially in light of recent observations that the mRNA for certain RTKs are overexpressed in some cancers (5).

The methods developed in the present invention for detecting mRNA expressed by the kdp gene can be readily used by those of ordinary skill in the art for the detection of mRNA species related to the kdp gene in any cell type and from any species. For this reason, the present invention embodies all mRNA segments which are the result of transcription of the kdp gene.

The present invention relates to methods for expression of the receptor protein, for example, in CMT-3 cells of monkey kidney origin. The receptor

- 6 -

protein, portions thereof, and mutated forms of the receptor protein may be expressed in many other cells by those skilled in the art using methods similar to those described in this application. For this reason, 5 the present invention embodies all proteins encoded by the human KDR gene and proteins encoded by related genes found in other species.

The present invention further relates to methods for studying the interaction of VEGF to the 10 expressed KDR protein. Recent work in the literature (17) indicates that VEGF is one member of a family of related proteins, and the interaction of growth factors similar to VEGF with the KDR protein can readily be studied by those skilled in the art using methods 15 similar to those described in this application. These methods can readily be modified to study the interaction of candidate pharmaceuticals with the KDR protein towards the goal of developing an antagonist or agonist of VEGF action. For this reason, the present 20 invention embodies methods for studying the interaction of VEGF and VEGF-related growth factors with the KDR protein.

25

#### BRIEF DESCRIPTION OF THE DRAWINGS

30

Figure 1 depicts a schematic representation of three receptor tyrosine kinase subclasses (6). KI is kinase insert domain; PTK is kinase domain; cys is cysteine rich region.

35

Figure 2 depicts the two sets of primers used for PCR (SEQ ID No: 1 and 2). The nucleotide sequences in appropriate regions of the four known type III receptor tyrosine kinase cDNAs are aligned and degenerate oligonucleotide primers are designed based upon the consensus sequences.

- 7 -

Figure 3 depicts the amplification of the kinase insert domains using PCR. DNA segments encoding the kinase insert domains of type III receptor tyrosine kinases are amplified by PCR. A sample (5  $\mu$ l) is run on a 1.0% agarose gel which is stained with ethidium bromide. DNA size standards (123 bp ladder; Bethesda Research Laboratories, Bethesda, MD) are run as well.

5

10

15

Figure 4 depicts the DNA sequence of the two PCR products (Panel A: 363 bp segment derived from the 420 bp product (SEQ ID NO: 3); Panel B: 251 bp product (SEQ ID NO: 4)). The two products are purified by agarose gel electrophoresis, digested with Sall and EcoRI, and cloned into the plasmid vector pBlueScribe(+) (Stratagene; San Diego, CA). The 420 bp PCR product is digested to 363 bp during this procedure. The DNA sequences for the primers used in the amplification are underlined.

20

25

Figure 5A depicts a computer assisted comparison of the DNA sequence for the 363 bp DNA segment derived from the 420 bp PCR product with the sequence of a DNA segment of the PDGF receptor (SEQ ID NO: 5) (18). A region of strong homology between the 363 bp segment derived from the 420 bp PCR product and the PDGF receptor is contained in a box. Figure 5B depicts a computer assisted comparison of the DNA sequence for the 251 bp PCR product with the sequence of a DNA segment of the FGF receptor (SEQ ID NO: 6) (7).

30

35

Figure 6 depicts the strategy used for sequencing the insert portions of clones BTIII081.8 and BTIII129.5 and BTIV169. The sequencing reaction uses either synthetic oligonucleotides (represented by boxes at the start of an arrow), or the M13 universal primer (no box) to initiate the reaction. In some cases, portions of these DNA segments are isolated using the

- 8 -

restriction enzymes indicated in the figure, and  
subcloned back into the plasmid vector pUC118, so that  
the M13 universal primer can be used. The position of  
the stop codon in BTIII129.5 is indicated. The coding  
5 portions of these DNA segments are shown at the bottom  
of the figure. The relative positions of the 1)  
membrane spanning portion, 2) kinase domains, and 3)  
kinase insert domain are indicated. The position of  
these structural features within the KDR derived DNA  
10 segments is compared in relation to their position in  
the PDGF-receptor ("PDGF-R").

Figure 7 depicts the DNA and predicted amino  
acid sequence of KDR, plus the stop codon (nucleotides  
1-4071 of SEQ ID NO. 7). The sequence of the DNA  
15 segment amplified by PCR is underlined (nucleotides  
2749-3105 of SEQ ID NO. 7). Cysteine residues in the  
putative extracellular domain are circled. Potential  
N-linked glycosylation sites are indicated by an  
asterisk. The putative membrane spanning region is  
20 enclosed in a box (nucleotides 2293-2367 of SEQ ID NO.  
7).

Figure 8 depicts a hydropathy plot of the  
predicted amino acid sequence for the KDR protein.

Figure 9 depicts a comparison of the  
predicted amino acid sequence in the putative  
25 intracellular portion of the KDR protein to the ckit  
proto-oncogene (SEQ ID No: 8) (3), the CSF-1 receptor  
(SEQ ID NO: 9) (4), and the PDGF receptor (SEQ ID NO:  
10) (18). Exact matches are indicated by an asterisk.  
Gaps are introduced to achieve maximum alignment. The  
30 putative ATP recognition site is indicated by three  
asterisks.

Figure 10 depicts the identification of kdp  
receptor mRNA by Northern blot analysis. Five  
35 micrograms of bovine aortic endothelial cell polyA+ RNA

- 9 -

are used. A nick-translated [<sup>32</sup>P] CTP-labelled EcoRI/BamHI DNA segment (nucleotides 1510-2417 of SEQ ID NO. 7) is used as a probe. Autoradiography is for 36 hours.

5           Figure 11 depicts the kdp gene in human and mouse DNA by Southern blot analysis. A nick translated [<sup>32</sup>P]CTP-labelled EcoRI/BamHI DNA segment (nucleotides 1510-2417 of SEQ ID NO. 7) is used as the probe. The probe is hybridized to Southern blots containing EcoRI  
10          digested DNA from human (lane 1), mouse (lane 2), and human-mouse hybrid cells (19) (lanes 3 and 4). The DNA used in lane 3 lacks the kdp locus, while DNA used in lane 4 contains the kdp locus.

15          Figure 12 depicts a Western blot analysis of CMT-3 cells which express the KDR protein. Cells are transfected with either the pcDNA1tkpASP vector alone (lane 1) or with that vector modified to contain the KDR gene (lane 2).  $2 \times 10^5$  cells and 1 microgram of DNA are used for each transfection. Forty-eight hours later, Western blot analysis is performed on the samples using the anti-KDR.PS23 polyclonal antibody at a dilution of 1:1000. Detection of reacting proteins is performed using an ECL system (Amersham, Chicago, IL).

25          Figure 13 depicts the results of [<sup>125</sup>I] VEGF binding to CMT-3 cells which express the KDR protein. Cells are transfected with either the vector alone (bars 1 and 2) or with the vector containing the KDR gene (bars 3 and 4). Forty-eight hours later, the samples are washed with phosphate buffered saline (PBS), and incubated with serum-free media containing 50 pM [<sup>125</sup>I] VEGF (specific activity equal to 4,000 cpm per fmol), for 90 minutes. Nonradioactive VEGF, 5 nM, is added to some samples (bars 2 and 4) to define

- 10 -

specific binding sites. The samples are washed with ice cold PBS, and the cells are transferred to gamma-counting tubes using 0.1% lubrol.

Figure 14 depicts the results of affinity cross-linking of [<sup>125</sup>I] VEGF to CMT-3 cells which express the KDR protein. CMT-3 cells are transfected with either the vector alone (lane 1) or with the vector containing the KDR gene (lane 2). Forty-eight hours later, the cells are washed in PBS, and serum free media containing 200 pM [<sup>125</sup>I] VEGF is added. After 90 minutes at room temperature, an affinity cross-linker disuccinimidyl suberate, 0.5 mM, is added for 15 minutes. The samples are then prepared for SDS-PAGE autoradiography.

15

#### DETAILED DESCRIPTION OF THE INVENTION

The strategy used to discover the DNA segments for the novel type III RTK gene begins with the design of two degenerate oligonucleotide primers based upon their homology to specific regions of the kinase domains of known RTK genes (Fig. 2) (3,4,7,18). In one embodiment, the polymerase chain reaction is then used to amplify DNA segments from a human endothelial cell cDNA library (designated HL 10246). The cDNA products from this step are each cloned into a plasmid vector designated pBlueScribe+ (Stratagene, San Diego, CA) and sequenced. Oligonucleotide probes are designed from potentially interesting sequences in order to screen the cDNA library for more full length clones of the novel cDNA.

The strategy just described provides several novel elements: 1) the DNA sequences of the oligonucleotide primers used during PCR; 2) the DNA sequence of the products generated by the polymerase chain

35

- 11 -

reaction; and 3) the DNA sequence of the final cloned DNA segments. Each of these elements of the invention described in this application will now be discussed in detail.

5       Figure 2 shows the rationale for choosing the oligonucleotide primers used in the PCR. The primers are designed to allow for the PCR amplification of the kinase insert domain of type III RTK genes. In order to design the primers, the DNA sequences of known type III  
10      RTK genes are aligned in specific regions of their catalytic domains, and a consensus sequence is chosen. The regions of the catalytic domains chosen in designing the primers flank the kinase insert domains of the receptor genes.

15      Primer 1 (SEQ ID No: 1) is designed from a region of the kinase domain 5' to the kinase insert domain, and consists of a mixture of four different 21mers. Primer 2 (SEQ ID NO: 2) is designed from a region of the kinase domain 3' to the kinase insert domain, and consists of a mixture of sixteen different 29mers with one inosine, indicated in SEQ ID NO: 2 by "NN".

20      SalI and EcoRI restriction sites are included at the 5' end of primers 1 and 2, respectively, to facilitate the subcloning of the amplified PCR products into plasmid vectors. Those skilled in the art may use other restriction sites; other minor modifications in the protocol above permits the design of primers without the inclusion of restriction sites.

25      The selection of these specific primers constitutes a novel approach towards identifying novel type III RTK genes. It had previously been shown (10) that primers designed from DNA sequences common to all tyrosine kinases allows for the identification of novel proteins. The present invention is the first to

- 12 -

contemplate the use of PCR to specifically target type III RTK.

The protocol used for PCR is as follows:  
Human endothelial cell cDNA (designated HL10246) is  
5 denatured by boiling and submitted to 30 cycles of PCR  
using 1 nmol of both primers in a final volume of 100  
μl. The timing is 1.5 minutes at 92°C, 2 minutes at  
50°C, and 2 minutes at 74°C. DNA from 5μl of sample is  
10 separated on a 1% agarose gel and stained with ethidium  
bromide.

Figure 3 shows the results of the PCR  
amplification. Two DNA products, with sizes 251 bp  
(SEQ ID NO: 4) and 420 bp, are visible when a sample of  
15 the reaction is electrophoresed on a 1.0% agarose gel  
and stained with ethidium bromide. The sizes of the  
two products are within the range expected for type III  
RTK genes (products derived from the FGF and PDGF  
receptor genes, which have the smallest and largest  
known kinase insert domains, would be 230 and 510 bp,  
20 respectively (20, 21)).

The DNA from four contiguous lanes with  
sizes ranging from 200 to 600 bp is electrophoresed  
onto DEAE filter paper, eluted from the paper with  
salt, and ethanol precipitated. The samples are  
25 incubated with 5 units of EcoRI and SalI. The  
restriction enzymes digest the 420 bp DNA segment to a  
363 bp DNA segment (SEQ ID NO: 3), due to the presence  
of an EcoRI site within the 420 bp DNA segment  
(nucleotide 2749, SEQ ID NO. 7). The restriction  
30 enzyme digested PCR products are then subcloned into  
the plasmid vector pBlueScribe(+). The recombinant  
clones are analyzed by sequencing using the  
dideoxy-method (22) using a United States Biochemical  
(Cleveland, Ohio) Sequenase Version 2.0 sequencing kit.  
35 Figure 4 shows the DNA sequences for the 251 bp PCR

- 13 -

product and the 363 bp DNA segment derived from the 420 bp PCR product.

Computer assisted comparison of the DNA sequence for the 363 bp segment of the 420 bp PCR product to databases of known DNA sequences reveals that the sequence is novel, because it shares strong sequence identity with the flanking catalytic domain of known type III RTK genes, but not their kinase insert domains. Figure 5A compares the DNA sequence for the 363 DNA segment with that for the PDGF receptor gene (SEQ ID NO: 5). Similar results are obtained using other type III RTK genes.

DNA sequencing of the 251 bp PCR product reveals a novel sequence containing both primers used for the amplification, but the sequence shows little homology to known tyrosine kinases. This is depicted in Figure 5B, which compares the DNA sequence for the 251 bp DNA segment with that for the FGF receptor (SEQ ID NO: 6). For this reason, further analysis of Product 1 is not pursued.

The protocols used during the PCR do not allow for amplification of the kinase insert domains of known receptor tyrosine kinases in the endothelial cell library used because of the low copy number of the message present in the library. There have been many studies on the effect of FGF on endothelial cell function (23,24) although there is evidence that the expression of the FGF receptor is developmentally regulated (7) and it is likely that the library used contains little or no cDNA for the FGF receptor.

An oligonucleotide probe, designed from the DNA sequence of the 363 bp segment, is synthesized (using an ABI 380 DNA Synthesizer) in order to screen the human endothelial cell cDNA library (HL10246) for the isolation of more full length clones containing the

- 14 -

363 bp DNA segment. The probe sequence is chosen from the region of the 363 bp DNA segment which shares little sequence homology with known RTK.

The screening of the endothelial cell cDNA library is conducted as follows: Lambda gt11 phage, 5  $\times 10^6$ , are adsorbed to *E. coli* LE392 for 15 minutes at 37°C prior to plating onto agar plates at a density of 5  $\times 10^5$  phage per plate. After allowing the phage plaques to develop at 37°C, plaque lifts are made using nitrocellulose filters, denatured in 0.4 N NaCl for 1 minute, and neutralized in 0.5 M Tris.HCl, pH 7.3, plus 1.5 M NaCl. The filters are washed with 2 x standard saline citrate (SSC) and then baked for 1.5 hour in a vacuum oven at 80°C. The filters are probed with an [32P] ATP end labeled synthetic oligonucleotide, 5' - 10 [32P] ATP end labeled synthetic oligonucleotide, 5' - TTTCCCTTGACGGAATCGTGCCCTTGTTGGT-3', which is the reverse complement of a DNA sequence contained in the PCR amplified product (Fig. 3). Hybridization is performed at 50°C in 5 x SSPE (167 mM NaCl, 10 mM sodium phosphate, pH 7.4, 1 mM EDTA), 2.5 x Denhardts, 0.5% sodium dodecyl sulfate (SDS), 100 µg/ml salmon sperm DNA. The filters are washed twice, 20 minutes per wash, with 2 x SSC plus 0.1% SDS at room temperature, followed by washing twice at 50°C with 0.1 X SSC plus 15 0.1% SDS; 20 minutes per wash. Positive clones are identified, picked and plaque purified.

20

25

30 Forty-five positive clones are obtained. Three of these positive clones are plaque purified and their phage DNA isolated. Digestion of the DNA with EcoRI and electrophoresis in agarose indicates that one clone, designated BTIII081.8, contains the largest insert, and subsequent analysis indicates that the DNA insert of this clone overlaps that of the inserts contained in other two purified clones (designated 35 BTIII079.11 and BTIII079.47A).

- 15 -

Digestion of the purified phage DNA of the clone designated BTIII081.8 with EcoRI results in DNA segments of 250 bp, 600 bp, and 1000 bp. Each of these three products is subcloned into the plasmid vector pUC118 and sequenced (Figure 6 shows the strategy used for sequencing). The orientation of the three fragments is determined by subcloning from the insert a BglII/BglII fragment into pUC118 and sequencing across the EcoRI junctions using a synthetic oligonucleotide to prime the sequencing reaction.

A restriction map is determined for each fragment (Figure 6). Various restriction site pieces are removed from the plasmids and recloned into pUC118 so that sequencing the resulting plasmids with the universal primer allows for sequencing most of the entire original fragments in both directions. Three oligonucleotide primers are required to sequence the entire cDNA in both directions. For the purposes of this application, this insert contains nucleotides numbered 1510-3406 (SEQ ID NO. 7).

A [<sup>32</sup>P]CTP-labelled, nick-translated EcoRI-BamHI DNA segment derived from clone BTIII081.8 (nucleotides 1510-2417 of SEQ ID NO. 7) is used as a probe to rescreen the original endothelial cell cDNA library for more 5' full length DNA segments of the gene from which the insert portion of BTIII081.8 is derived. The protocols used to isolate the overlapping clones are identical to that used to isolate BTIII081.8.

A synthetic oligonucleotide probe is designed with 29 nucleotides corresponding to part of the DNA sequence of the insert portion of the clone BTIII081.8 (nucleotides 3297-3325 of SEQ ID NO. 7) in order to rescreen the original endothelial cell cDNA library for more full 3' length DNA segments of the gene from which

- 16 -

the insert portion of BTIII081.8 is derived. The protocols used to isolate the overlapping clones are identical to that used to isolate BTIII081.8. Several positive clones for each of the 5' and 3' ends are identified and plaque purified.

5 One of the clones is designated BTIII200.2. The DNA from BTIII200.2 contains a 3.4 kb insert as determined by EcoRI digestion of the isolated phage DNA. EcoRI digestion of BTIII200.2 results in three 10 DNA fragments. One of these fragments (2.5 kb) is cloned into pUC119 and is designated BTIV006. The clone BTIV006 contains nucleotides numbered 7-2482. As described below, BTIV006 plus nucleotides 1-6 is 15 designated BTIV169. DNA sequencing of the 2.5 kb DNA insert (BTIV169) indicates that it overlaps over one thousand nucleotides of the DNA sequence of the insert portion of the clone BTIII081.8 (Figure 6) at the 5' end.

20 A second clone isolated from the cDNA library is designated BTIII129.5. The DNA from BTIII129.5 contains a 2.2 kb insert as determined by EcoRI digestion of the isolated phage DNA. DNA sequencing of the 2.2 kb DNA insert indicates that it overlaps over 25 five hundred nucleotides of the DNA sequence of the insert portion of the clone BTIII081.8 (Figure 6). The clone BTIII129.5 contains nucleotides numbered 2848-4236 (SEQ ID NO. 7). The DNA sequence for BTIII129.5 contains the stop codon TAA, defining the position of the 3' end of an open reading frame for the novel gene. Except for the first six nucleotides of the gene which are discussed below, these three clones define a gene encoding a growth factor receptor. These 30 three clones define a 4,062 nucleotide sequence of the open reading frame of the gene extending to the 3' end, followed by a 168 nucleotide non-coding region (SEQ ID 35

- 17 -

NO. 7). A sample of a lambda gt11 phage harboring the clone BTIII081.8 has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A., and has been assigned ATCC accession number 40,931. A sample of a lambda gt11 phage harboring the clone BTIII129.5 has been deposited with the American Type Culture Collection and has been assigned ATCC accession number 40,975. For reasons discussed below, a sample of the 5 clone BTIV006 was not deposited.

10

The aforementioned DNA segments (BTIII081.8, BTIII129.5, and BTIII200.2 (or BTIV006) encode 4062 nucleotides of the coding portion of a novel gene. The cDNA clones are incomplete in that a transcription 15 initiation coding for methionine is missing. After the isolation of these clones, Matthews et al. (25) reported the cloning of a gene homologue of KDR in mouse, which was referred to as Elk-1. Analysis of the nucleic acid and amino acid sequence of Elk-1 indicated that the addition of six nucleotides to the 5' end of 20 the isolated KDR clones would provide for a complete coding region.

To achieve this, an EcoRI-BamHI restriction fragment of BTIV200.2 is cloned into the plasmid pBlueScript KS (Stratagene, La Jolla, CA). The 5' end 25 of the inserted DNA is blunt ended with Klenow polymerase and Mung Bean nuclease. Next, the synthetic oligonucleotide TCGACGGCG ATG GAG (SEQ ID NO. 11) is cloned into this vector. The oligonucleotide contains 30 the sequence ATG GAG in frame with the downstream DNA insert. These nucleotides (ATG GAG) encode the amino acids methionine and glutamic acid, the first two amino acids encoded by the KDR gene. The resulting plasmid vector is designated BTIV140. This plasmid is purified 35 on a CsCl gradient.

- 18 -

The purified plasmid is designated BTIV169. The insert of BTIV169 contains nucleotides 1-2400 (SEQ ID NO. 7) of the KDR gene. A sample of the plasmid pBlueScript KS which contains the clone BTIV169 has been deposited with the American Type Culture Collection and has been assigned ATCC accession number 75200.

Thus, together the clones BTIII081.8, BTIII129.5 and BTIV169 comprise the entire open reading frame of 4,068 nucleotides for the novel KDR gene. As will be discussed below, the KDR gene expresses the novel KDR receptor which binds specifically to the growth factor VEGF.

DNA sequencing of BTIII081.8, BTIII129.5 and BTIV169 (SEQ ID NO. 7) shows that the newly isolated gene is similar to, but distinct from, previously identified type III RTK. The predicted amino acid sequence (SEQ ID NO. 7) contains several structural features which demonstrate that the novel gene is a type III RTK. These structural features are summarized as follows:

1) A hydrophathy plot of the predicted amino acid sequence indicates a single membrane spanning region (see Figure 8). This is characteristic of a type III RTK (Figure 7).

2) The putative amino-terminal 762 amino acid portion of the receptor has structural features of extracellular receptor ligand binding domains (1), including regularly spaced cysteines and 18 potential N-linked glycosylation sites (Figure 7).

3) The predicted amino acid sequence of the carboxy-terminal 530 amino acid portion contains an ATP-binding site at lysine 868, 22 amino acids downstream from the consensus ATP recognition sequence Gly-X-Gly-X-X-Gly (26) (Figure 8).

- 19 -

4) Within the kinase domain there is a  
55-60% identical match in amino acid sequence to three  
other type III receptor tyrosine kinases: skit  
proto-oncogene (SEQ ID NO: 8), CSF-1 (SEQ ID NO: 9) and  
5 PDGF (SEQ ID NO: 10) (Figure 9).

5) The predicted kinase domain contains a  
kinase insert domain of approximately 71 amino acids.  
As indicated in Figure 9, this portion of the amino  
acid sequence shares little sequence homology with  
10 other type III RTK.

The endothelial cell library can be further  
screened to isolate the 5' untranslated region and  
genomic clones can be generated so as to isolate the  
promoter region for the KDR gene.

15 In addition to the DNA sequence described for  
the KDR gene (SEQ ID NO. 7), the present invention  
further comprises DNA sequences which, by virtue of the  
redundancy of the genetic code, are biologically  
equivalent to the sequences which encode for the  
20 receptor, that is, these other DNA sequences are  
characterized by nucleotide sequences which differ from  
those set forth herein, but which encode a receptor  
having the same amino acid sequences as those encoded  
by the DNA sequences set forth herein.

25 In particular, the invention contemplates  
those DNA sequences which are sufficiently duplicative  
of the sequence of SEQ ID NO. 7 so as to permit  
hybridization therewith under standard high stringency  
Southern hybridization conditions, such as those  
30 described in Sambrook et al. (27), as well as the  
biologically active proteins produced thereby.

This invention also comprises DNA sequences  
which encode amino acid sequences which differ from  
those of the novel receptor, but which are the  
35 biological equivalent to those described for the

- 20 -

receptor. Such amino acid sequences may be said to be biologically equivalent to those of the receptor if their sequences differ only by minor deletions from or conservative substitutions to the receptor sequence, such that the tertiary configurations of the sequences are essentially unchanged from those of the receptor.

For example, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, as well as changes based on similarities of residues in their hydropathic index, can also be expected to produce a biologically equivalent product. Nucleotide changes which result in alteration of the N-terminal or C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein. It may also be desirable to eliminate one or more of the cysteines present in the sequence, as the presence of cysteines may result in the undesirable formation of multimers when the protein is produced recombinantly, thereby complicating the purification and crystallization processes. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Therefore, where the terms "KDR gene" or "KDR protein" are used in either the specification or the claims, each will be understood to encompass all such modifications and variations which result in the production of a biologically equivalent protein.

- 21 -

In addition to the full length gene and protein, the invention encompasses biologically active fragments of each. By "biologically active" is meant a protein fragment which qualitatively retains the receptor activity of the larger KDR protein, or, in the case of a nucleotide sequence, which encodes such a protein fragment. It also refers, for purposes of antibody production, to fragments which are capable of eliciting production of antibodies capable of binding to the receptor protein.

To determine the size of the mRNA transcribed from the kdp gene, Northern blot hybridization experiments are carried out using an EcoRI/BamHI DNA segment (nucleotides 1510-2417, SEQ ID NO. 7) as a hybridization probe. The DNA used for the probe does not contain any portion of the putative kinase domain, and shares little sequence homology to other tyrosine kinases. The Northern blot analysis (Figure 10) shows that a 7 kb band is visualized in cytoplasmic poly(A)+ RNA of ABAE bovine aortic endothelial cells. This transcript differs in size from previously reported transcripts for known type III RTK (7,18).

The isolated cDNA is significant for several reasons. The cDNA encodes a novel type III receptor tyrosine kinase. The homology between the sequence of this cDNA and that of other receptors, as well as structural properties implied by the predicted amino acid sequence confirm the relationship. Receptors for growth factors should have tremendous utility in drug development as they face the outside of the cell and thus are among the best targets for drugs. In addition, the cellular levels of some receptors, in particular the neu proto-oncogene, increase during some cancers. This has been taken advantage of in designing diagnostic tests for these cancers.

- 22 -

Southern analysis demonstrates that the kdp gene is present in mouse as well as human DNA. Mouse and human (HeLa cell) DNA, 15 µg of each, are digested with 10 units of EcoRI and electrophoresed on a 0.7% agarose gel. The DNA is transferred onto nitrocellulose. The filter is hybridized to a [<sup>32</sup>P]CTP-labelled cDNA probe made by nick translating an EcoRI/BamHI fragment from the 5' end of the kdp cDNA (nucleotides 1510-2417, SEQ ID NO. 7). Hybridization is conducted at 30°C in 5 X SSPE, 50% formamide, 0.1% SDS, plus 150 µg/ml salmon sperm DNA. The DNA probe hybridizes to Southern blots containing EcoRI digested DNA. After 48 hours, the filter is washed at room temperature in 2 X SSC plus 0.1% SDS for 20 minutes, followed by two 20 minute washes at 40°C with 0.1 X SSC plus 0.1% SDS. Autoradiography is then performed for 48 hours. As shown in Figure 11, radioactively labelled DNA is present in both human and mouse samples. This indicates that the kdp gene is present in both species.

An experiment is conducted to ascertain the genetic locus of kdp on human chromosomes. Thirty-eight cell hybrids from 18 unrelated human cell lines and four mouse cell lines are examined (19). A DNA probe hybridizes to Southern blots which contain EcoRI digested DNA from the human-mouse hybrids (using the procedure and DNA probe for human and mouse tissue described in relation to Figure 11). Table I sets forth the results of the segregation of kdp with human chromosomes in EcoRI digested human-mouse somatic cell hybrid DNA:

- 23 -

Table I

5	Chromosome	Concordant #		Discordant #		% Discordancy
		(+/+)	(-/-)	(+/-)	(-/+)	
10	1	4	19	8	4	34
	2	8	18	5	6	30
	3	11	12	3	9	34
	4	14	24	0	0	0
	5	7	14	7	10	45
	6	7	19	7	5	32
	7	11	14	3	8	31
15	8	8	11	6	13	50
	9	3	20	10	4	38
	10	12	9	2	14	43
	11	9	13	4	11	41
	12	9	10	5	14	50
20	13	7	18	7	6	34
	14	11	8	3	16	50
	15	9	15	5	8	35
	16	7	19	7	5	32
	17	12	7	2	16	49
25	18	11	14	3	10	34
	19	7	18	7	6	34
	20	9	10	5	14	50
	21	11	9	3	15	47
	22	3	16	10	7	47
30	X	8	10	3	8	38

30

The scoring is determined by the presence (+) or absence (-) of human bands in the hybrids on Southern blots prepared in a similar to those shown in Figure 11. The scoring is compared to the presence or absence of human chromosomes in each hybrid. A 0%

35

- 24 -

discordancy indicates a matched segregation of the DNA probe with a chromosome. Three fragments, approximately 6.5 kb, 3.1 kb, and 0.7 kb in size are detected in digests of human DNA (Figure 11), and in all hybrids which had retained human chromosome 4 (Table I). All other chromosomes are excluded in at least 11 discordant hybrids (Table I). The results of Figure 11 and Table I demonstrate that the genetic locus of kdp is on human chromosome 4.

It is noteworthy that both the ckit (3) and the type A PDGF (28) receptor genes map to human chromosome 4. The finding that the genetic locus of kdp is on human chromosome 4 provides further evidence that the novel receptor of this invention is a type III receptor tyrosine kinase.

The next step after identifying the entire coding portion of the kdp gene is to express the receptor protein encoded by that gene. The receptor protein is then utilized so as to identify the growth factor which binds specifically to the receptor.

The receptor protein is expressed using established recombinant DNA methods. Suitable host organisms include bacteria, viruses, yeast, insect or mammalian cell lines, as well as other conventional organisms. For example, CMT-3 monkey kidney cells are transfected with a vector containing the complete coding region of the KDR gene.

The complete coding portion of the KDR gene is assembled by sequentially cloning into pUC119 three DNA fragments derived from BTIII081.8, BTIII129.5, and BTIV169. First, a SmaI-EcoRI fragment of clone BTIII129.5 (nucleotides 3152-4236, SEQ ID NO. 7) is blunt ended with Klenow polymerase and introduced into a SmaI site in pUC119. Next, a BamHI-SmaI fragment of clone BTIII081.8 (nucleotides 2418-3151, SEQ ID NO. 7)

- 25 -

is introduced at a BamHI-SmaI site. Finally, a SalI-BamHI fragment of clone BTIV169 (nucleotides 1-2417, SEQ ID NO. 7) is introduced at a SalI-BamHI site. Part of the cloning site of pUC119 is contained 5 in the SalI-BamHI fragment, 5' to the KDR gene. In order to clone the complete coding portion into an expression vector, the assembled DNA (in pUC119) is digested with SalI and Asp118 and recloned into the eukaryotic expression vector pcDNA1tkpASP.

10 This vector is a modification of the vector pCDNA1 (Invitrogen; San Diego, CA). Specifically, the ampicillin resistance gene is cloned from pBR322 into pCDNA1. A small SV40 T splice and the SV40 polyadenylation signal are then removed and are replaced with a Herpes Simplex Virus-1 polyadenylation signal.  
15 Finally, a cytomegalovirus intermediate early splice is inserted 5' to the cloning site to yield pcDNA1tkpASP.

20 Transfection of CMT-3 cells is done using DEAE-dextran. Forty-eight hours after transfection, expression of the novel receptor is monitored using Western blot analysis as follows.

25 An antibody is used to assay the expressed receptor protein. The predicted amino acid sequence of the receptor is used to generate peptide-derived antibodies to the receptor by conventional techniques. The presence of the novel receptor protein is confirmed by Western blot hybridization.

30 Specifically, a synthetic peptide with 13 residues is synthesized based on the 12 residues corresponding to amino acids 986-997 of the putative amino acid sequence of the KDR protein (SEQ ID NO. 7), with a cysteine residue linked to the lysine (amino acid 997). The cysteine facilitates coupling of the peptide to a macromolecule which functions as a carrier 35 for the peptide. For example, the peptide is coupled

- 26 -

5 to keyhole limpet haemocyanin (KLH) using  $\text{n}$ -maleimidobenzoyl-N-hydroxysuccinimide ester. Other conventional carriers may be used such as human and bovine serum albumins, myoglobins,  $\beta$ -galactosidase, penicillinase and bacterial toxoids, as well as synthetic molecules such as multi-poly-DL-alanyl-poly-L-lysine and poly-L-lysine.

10 Rabbits are immunized with the peptide-KLH conjugate to raise polyclonal antibodies. After different periods of time, serum is collected from the rabbits. The IgG fraction of the serum is then purified using a protein A Sepharose column (Pharmacia LKB, Uppsala, Sweden) to obtain the antibody which is designated anti-KDR.PS23.

15 A sample of the expressed KDR protein is subjected to SDS-PAGE using a 7% acrylamide gel under standard conditions. The protein band is then transferred on to nitrocellulose paper for Western blot analysis and the anti-KDR.PS23 antibody is added at a dilution of 1:1,000 to allow the antibody to react with the protein present. A second antibody, goat 20 anti-rabbit antibody to rabbit IgG, which binds to anti-KDR.PS23, is then added. The detection of proteins which react with the antibodies is performed by autoradiography of bands using an ECL system 25 (Amersham, Chicago, IL). The results are depicted in Figure 12.

30 Figure 12 shows that a 190 kD protein is present in the cells transfected with the vector containing the KDR gene, but is absent in cells transfected with vector alone. The size of this 35 protein is consistent with it being encoded by the KDR gene, in that the predicted amino acid sequence for the unglycosylated KDR protein is 156 kD, and that sequence contains 18 putative extracellular glycosylation sites

- 27 -

which would account for the balance of the size seen in the 190 kD band.

The expressed receptor is then used to identify the growth factor which interacts with the receptor. In order to test the hypothesis that the KDR protein is a receptor for VEGF, radioligand binding studies are performed. VEGF (provided by D. Gospodarowicz) is radiolabelled with  $^{125}\text{I}$ . Cells are transfected with either the vector pcDNA1tkpASP alone (bars 1 and 2 of Figure 13) or with the vector containing the KDR gene (bars 3 and 4). Forty-eight hours later, the transfected cell samples are washed with PBS and then incubated for 90 minutes with serum-free media containing 50 pM [ $^{125}\text{I}$ ]VEGF (specific activity equal to 4,000 cpm per fmol). Excess nonradioactive VEGF, 5 nM, is added to some samples (bars 2 and 4) to define specific binding sites. The samples are washed with ice cold PBS, and the cells are transferred to gamma-counting tubes using a detergent, 0.1% lubrol.

The results of the radioligand binding studies are depicted in Figure 13. Figure 13 shows that CMT-3 cells transfected with vector containing the KDR gene contain specific binding sites for [ $^{125}\text{I}$ ]VEGF (compare bars 3 and 4), while cells transfected with vector alone do not (compare bars 1 and 2).

Further evidence that the KDR gene encodes a receptor for VEGF is demonstrated by affinity cross-linking studies (Figure 14). Figure 14 depicts the results of affinity cross-linking of [ $^{125}\text{I}$ ]VEGF to CMT-3 cells which express the KDR protein. CMT-3 cells are transfected with either the pcDNA1tkpASP vector alone (lane 1 of Figure 14) or with the vector containing the KDR gene (lane 2). Forty-eight hours later, the cells are washed in PBS, and serum free

- 28 -

5 media containing 200 pM [<sup>125</sup>I]VEGF is added. After 90 minutes at room temperature, an affinity cross-linker disuccinimidyl suberate (Pierce Biochemicals, Rockford, IL), 0.5mM, is added for 15 minutes. The samples are then subjected to SDS-PAGE autoradiography.

10 Three protein bands are seen in SDS-PAGE autoradiograms from samples of CMT-3 cells transfected with the KDR gene and cross-linked to [<sup>125</sup>I]VEGF (lane 1). The size of band 1 (235 kD) is consistent with it being the 190 kD protein seen by Western blot analysis (Figure 12), because a 45 kD [<sup>125</sup>I] VEGF dimer plus 190 kD would migrate in a manner identical to band 1. The origin of band 2 is not clear, but may represent an altered glycosylation form of band 1. Band 3 (22.5 kD) is most likely VEGF itself, and can be seen faintly in cells transfected with vector alone (lane 2).

15

20 The novel KDR gene of this invention is significant for several reasons. Studies of the cellular mechanisms by which receptors function in signal transduction have led in the past to a better understanding of how cells grow in both normal and diseased states. Receptor tyrosine kinases, in particular, have received a great deal of attention because of the observation that a number of RTK are the cellular counterparts for viral oncogenes, implying a direct correlation between changes in the expression of RTK and cancer. In view of this, it is likely that pharmaceuticals targeted at RTK will inhibit the changes in cell growth associated with cancer. In addition, it is likely that monitoring the levels of expression of RTK will prove valuable in diagnosing the onset of cancer.

25

30 The described cDNA is isolated from a human endothelial cell library. Endothelial cells participate in angiogenesis, the formation of new blood

35

- 29 -

capillaries. Previous work directed towards identifying the growth factors which regulate angiogenesis have primarily focused upon FGF (13), although recent evidence has indicated that other 5 growth factors may be involved as well (12,15,29). This evidence consists of the observations that: 1) FGF does not contain a signal sequence (24) and thus may not be secreted from cells in a manner consistent with the tight regulation of angiogenesis, and 2) 10 endothelial cells synthesize FGF and yet are normally resting (15). Our discovery, then, of a novel growth factor receptor may ultimately clarify these inconsistencies and lead to a better understanding of endothelial cell function.

15 The teachings of this invention can be readily used by those skilled the art for the purpose of testing pharmaceuticals targeted at the KDR protein. Two examples of approaches which can be used for this purpose are now given.

20 First, the methods described in this invention for studying the interaction of VEGF with KDR protein can be used to test for pharmaceuticals which will antagonize that interaction. For these studies, cells expressing the KDR protein are incubated with 25 [<sup>125</sup>I]VEGF, together with a candidate pharmaceutical. Inhibition of radioligand binding is tested for; significant inhibition indicates the candidate is an antagonist. Permanent expression of the KDR protein in a cell type such as NIH3T3 cells would make these 30 studies less laborious. This can be easily achieved by those skilled in the art using the described methods.

35 Second, using the teachings of this invention, those skilled in the art can study structural properties of the KDR protein involved in receptor function. This structural information can

- 30 -

then be used to more rationally design pharmaceuticals which inhibit that function. Mutagenesis of the KDR gene by well established protocols is one approach, crystallization of the receptor binding site is another.

5

10

15

20

25

30

35

- 31 -

Bibliography

1. Yarden Y., and A. Ullrich, Ann. Rev. Biochem., 57, 433-478 (1988).
- 5 2. Bargmann, C., et al., Nature, 319, 226-230 (1986).
- 10 3. Yarden, Y., et al., EMBO J., 6, 3341-3351 (1987).
4. Coussens, L., et al., Nature, 320, 277-280 (1986).
- 15 5. Slamon, D., et al., Science, 244, 707-712 (1989).
6. Ullrich, A. and Schlessinger, J., Cell, 61, 203-212 (1990).
- 15 7. Ruta, M., et al., Oncogene, 3, 9-15 (1988).
- 20 8. Strathmann, M., et al., Proc. Natl. Acad. Sci., 86, 8698-8702 (1989).
9. Streuli, M., et al., Proc. Natl. Acad. Sci., 86, 8698-8702 (1989).
- 20 10. Wilkes, A.F., Proc. Natl. Acad. Sci., 86, 1603-1607 (1989).
11. Folkman, J., and Klagsbrun, M., Science, 235, 442-445 (1987).
- 25 12. Ishikawa, F., et al., Nature, 338, 557-562 (1989).
13. Baird, A., and Bohlen, P., in Peptide Growth Factors and Their Receptors, pages 369-418 (Spron, M.B., and Roberts, A.B., eds. 1990).
- 30 14. Senger, D.R., et al., Science, 219, 983-985 (1983).
15. Gospodarowicz, D., et al., Proc. Natl. Acad. Sci., 86, 7311-7315 (1989).
- 35 16. Leung, D.W., et al., Science, 246, 1306-1309 (1989).

- 32 -

17. Maglione, D., et al., Proc. Natl. Acad. Sci., 88, 9267-9271 (1991).
18. Gronwald, R., et al., Proc. Natl. Acad. Sci., 85, 3435-3439 (1988).
- 5 19. Shows, T., et al., Somat. Del. Mol. Gen., 10, 315-318 (1984).
20. Rainer, G., et al., Proc. Natl. Acad. Sci., 85, 3435-3439 (1988).
- 10 21. Lee, P. L., et al., Science, 245, 57-60 (1989).
22. Sanger, F., et al., Proc. Natl. Acad. Sci., 74, 5463-5467 (1977).
- 15 23. Folkman, J., Cancer Res., 46, 467-473 (1986).
24. Burgess, W. and Maciag, T., Ann. Rev. Biochem., 58, 575-606 (1989).
- 15 25. Matthews, W., et al., Proc. Natl. Acad. Sci., 88, 9026-9030 (1991).
- 20 26. Hannink, M. and Donoghue, D., Proc. Natl. Acad. Sci., 82, 7894-7898 (1985).
27. Sambrook, J., et al., Molecule Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).
- 25 28. Matsui, T., et al., Science, 243, 800-804 (1989).
29. Conn, G., et al., Proc. Natl. Acad. Sci., 87, 2628-2632 (1990).

30

35

- 33 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5 (i) APPLICANT: Terman, Bruce I  
Carrion, Miguel E

(ii) TITLE OF INVENTION: Identification of a  
Novel Human Growth Factor Receptor

10 (iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Alan M. Gordon  
American Cyanamid Company

(B) STREET: 1937 West Main Street, P.O. Box 60

20 (C) CITY: Stamford

(D) STATE: Connecticut

(E) COUNTRY: USA

25 (F) ZIP: 06904

(v) COMPUTER READABLE FORM:

30 (A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC AT

(C) OPERATING SYSTEM: MS-DOS

- 34 -

(D) SOFTWARE: ASCII from IBM DW 4

(vi) CURRENT APPLICATION DATA:

5 (A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

10 (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/657,236

15 (B) FILING DATE: February 22, 1991

(viii) ATTORNEY/AGENT INFORMATION:

20 (A) NAME: Gordon, Alan M.

20 (B) REGISTRATION NUMBER: 30,637

(C) REFERENCE/DOCKET NUMBER: 31,298-01

25 (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 203 321 2719

(B) TELEFAX: 203 321 2971

30 (C) TELEX:

(2) INFORMATION FOR SEQ ID NO: 1:

35 (i) SEQUENCE CHARACTERISTICS:

- 35 -

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

5

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTCGAC AAY CTG TTG GGR GCC TGC AAC 27

15

(2) INFORMATION FOR SEQ ID NO: 2 :

(i) SEQUENCE CHARACTERISTICS:

20

(A) LENGTH: 35 base pairs

(B) TYPE: nucleic acid

25

(C) STRANDEDNESSS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GAATTC AG CAC KTT NCT RGC YGC CAG GTC TGY GTC 35

35

- 36 -

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

5 (A) LENGTH: 363 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAA TTC TGC AAA TTT GGA AAC CTG TCC ACT TAC CTG 36

AGG ACG AAG AGA AAT GAA TTT GTC CCC TAC AAG ACC 72

20 AAA GGG GCA CGA TTC CGT CAA GGG AAA GAC TAC GTT 108

GGA GCA ATC CCT GTG GAT CTG AAA CGG CGC TTG GAC 144

25 ACG CAT CAC CAG TAG CCA GAG CTC AGC CAG CTC TGG 180

ATT TGT GGA GGA GAA GTC CCT CAG TGA TGT AGA AGA 216

AGA GGA AGC TCC TGA AGA TCT GTA TAA GGA CTT CCT 252

30 GAC CTT GGA GCA TCT CAT CTG TTA CAG TTT CCA AGT 288

GGC TAA GGG CAT GGA GTT CTT GGC ATC GCG AAA GTG 324

35 TAT CCA CAG AGA CCT GGC AGC CAG GAA CGT GCT GAA 360

- 37 -

TTC

363

## (2) INFORMATION FOR SEQ ID NO: 4 :

5

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251 base pairs

10

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

20	GTC GAC AAT CTG TTG GGG GCC TGC ACC ATC CCA ACA	36
	TCC TGC TGC TCT ACA ACT ATT TTT ATG ACC GGA GGA	72
	GGA TCT ACT TGA TTC TAG AGT ATG CCC CCC GCG GAG	108
25	CTC TAC AAG GAG CTG CAG AAG AGC TGC ACA TTT GAC	144
	GAG CAG CGA ACA GCC ACG ATC ATG GAG GAG TTG GCA	180
30	GAT GCT CTA ATG TAC TGC CGT GGG AAG AAG GTG ATT	216
	CAC AGA GAC CTG GCA GCC AGC AAC GTG CTG AAT TC	251

35

(2) INFORMATION FOR SEQ ID NO: 5:

- 38 -

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 510 base pairs

5

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

10

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

15

(A) NAME/KEY: PDGF Receptor DNA

(B) LOCATION: Internal sequence

20

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Gronwald, R., et al.

(B) JOURNAL: Proc. Natl. Acad. Sci.

25

(C) VOLUME: 85

(D) PAGES: 3435-3439

30

(E) DATE: 1988

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AAC CTG TGG GGG CCT GCA CCA AAG GAG GAC CAT CTA 36

- 39 -

	TAT CAT CTA TAT CAT CAC TGA GTA CTG CCG CTA CGG	72
	AGA CCT GGT GGA CTA CCT GCA CCG CAA CAA ACA CAC	108
5	CTT CCT GCA GCA CCA CTC CGA CAA GCG CCG CCC GCC	144
	CAG CGC GGA GCT CTA CAG CAA TGC TCT GCC CGT TGG	180
	GCT CCC CCT GCC CAG CCA TGT GTC CTT GAC CGG GGG	216
10	AGA GCG ACG GTG GCT ACA TGG ACA TGA GCA AGG ACG	252
	AGT CGG TGG ACT ATG TGC CCA TGC TGG ACA TGA AAG	288
15	GAG ACG TCA AAT AGC AGA CAT CGA GTC CTC CAA CTA	324
	CAT GGC CCC TTA CGA TAA CTA CGT TCC CTC TGC CCC	360
	TGA GAG GAC CTG CCG AGC AAC TTT GAT CAA CGA GTC	396
20	TCC AGT GCT AAG CTA CAT GGA CCT CGT GGG CTT CAG	432
	CTA CCA GGT GGC CAA TGG CAT GGA GTT CTG GCC TCC	468
25	AAG AAC TGC GTC CAC AGA GAC CTG GCG GCT AGG AAC	504
	GTC CTT	510

30 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 255 base pairs

- 40 -

(B) TYPE: nucleic acid

(C) STRANDEDNESSS: single

5 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10 (ix) FEATURE:

10 (A) NAME/KEY: FGF Receptor DNA

(B) LOCATION: Internal sequence

15 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Ruta, M., et al.

(B) JOURNAL: Oncogene

20 (C) VOLUME: 3

(D) PAGES: 9-15

25 (E) DATE: 1988

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AAC CTG CTG GGG GCC TGC ACG CAG GAT GGT CCC TTG 36

30 TAT GTC ATC GTG GAG TAT GCC TCC AAG GGC AAC CTG 72

CGG GAG TAC CTG CAG ACC CGG AGG CCC CCA GGG CTG 108

35 GAA TAC TGC TAT AAC CCC AGC CAC AAC CCA GAG GAG 144

- 41 -

CAG CTC TCC TCC AAG GAC CTG GTG TCC TGC GCC TAC 180  
CAG GAG GCC CGA GGC ATG GAG TAT CTG GCC TCC AAG 216  
5 AAG TGC ATA CAC CGA GAC CTG GCA GCC AGG AAT GTC 252  
CTG 255

(2) INFORMATION FOR SEQ ID NO: 7:

10

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 4236 base pairs

20

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

25

ATG GAG AGC AAG GTG CTG CTG GCC GTC GCC CTG 33  
Met Glu Ser Lys Val Leu Leu Ala Val Ala Leu  
1 5 10

30

TGG CTC TGC GTG GAG ACC CGG GCC GCC TCT GTG GGT 69  
Trp Leu Cys Val Glu Thr Arg Ala Ala Ser Val Gly  
15 20

35

TTG CCT AGT GTT TCT CTT GAT CTG CCC AGG CTC AGC 105  
Leu Pro Ser Val Ser Leu Asp Leu Pro Arg Leu Ser  
25 30 35

- 42 -

	ATA CAA AAA GAC ATA CTT ACA ATT AAG GCT AAT ACA Ile Gln Lys Asp Ile Leu Thr Ile Lys Ala Asn Thr	141
	40	45
5	ACT CTT CAA ATT ACT TGC AGG GGA CAG AGG GAC TTG Thr Leu Gln Ile Thr Cys Arg Gly Gln Arg Asp Leu	177
	50	55
10	GAC TGG CTT TGG CCC AAT AAT CAG AGT GGC AGT GAG Asp Trp Leu Trp Pro Asn Asn Gln Ser Gly Ser Glu	213
	60	65
	65	70
15	CAA AGG GTG GAG GTG ACT GAG TGC AGC GAT GGC CTC Gln Arg Val Glu Val Thr Glu Cys Ser Asp Gly Leu	249
	75	80
20	TTC TGT AAG ACA CTC ACA ATT CCA AAA GTG ATC GGA Phe Cys Lys Thr Leu Thr Ile Pro Lys Val Ile Gly	285
	85	90
	90	95
	AAT GAC ACT GGA GCC TAC AAG TGC TTC TAC CGG GAA Asn Asp Thr Gly Ala Tyr Lys Cys Phe Tyr Arg Glu	321
	100	105
25	ACT GAC TTG GCC TCG GTC ATT TAT GTC TAT GTT CAA Thr Asp Leu Ala Ser Val Ile Tyr Val Tyr Val Gln	357
	110	115
30	GAT TAC AGA TCT CCA TTT ATT GCT TCT GTT AGT GAC Asp Tyr Arg Ser Pro Phe Ile Ala Ser Val Ser Asp	393
	120	125
	125	130
35	CAA CAT GGA GTC GTG TAC ATT ACT GAG AAC AAA AAC Gln His Gly Val Val Tyr Ile Thr Glu Asn Lys Asn	429
	135	140

- 43 -

	AAA ACT GTG GTG ATT CCA TGT CTC GGG TCC ATT TCA	465
	Lys Thr Val Val Ile Pro Cys Leu Gly Ser Ile Ser	
	145 150 155	
5	AAT CTC AAC GTG TCA CTT TGT GCA AGA TAC CCA GAA	501
	Asn Leu Asn Val Ser Leu Cys Ala Arg Tyr Pro Glu	
	160 165	
10	AAG AGA TTT GTT CCT GAT GGT AAC AGA ATT TCC TGG	537
	Lys Arg Phe Val Pro Asp Gly Asn Arg Ile Ser Trp	
	170 175	
15	GAC AGC AAG AAG GGC TTT ACT ATT CCC AGC TAC ATG	573
	Asp Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met	
	180 185 190	
20	ATC AGC TAT GCT GGC ATG GTC TTC TGT GAA GCA AAA	609
	Ile Ser Tyr Ala Gly Met Val Phe Cys Glu Ala Lys	
	195 200	
25	ATT AAT GAT GAA AGT TAC CAG TCT ATT ATG TAC ATA	645
	Ile Asn Asp Glu Ser Tyr Gln Ser Ile Met Tyr Ile	
	205 210 215	
30	GTT GTC GTT GTA GGG TAT AGG ATT TAT GAT GTG GTT	681
	Val Val Val Val Gly Tyr Arg Ile Tyr Asp Val Val	
	220 225	
35	CTG AGT CCG TCT CAT GGA ATT GAA CTA TCT GTT GGA	717
	Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly	
	230 235	
40	GAA AAG CTT GTC TTA AAT TGT ACA GCA AGA ACT GAA	753
	Glu Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu	
	240 245 250	

- 44 -

	CTA AAT GTG GGG ATT GAC TTC AAC TGG GAA TAC CCT	789
	Leu Asn Val Gly Ile Asp Phe Asn Trp Glu Tyr Pro	
	255	260
5	TCT TCG AAG CAT CAG CAT AAG AAA CTT GTA AAC CGA	825
	Ser Ser Lys His Gln His Lys Lys Leu Val Asn Arg	
	265	270
	275	
10	GAC CTA AAA ACC CAG TCT GGG AGT GAG ATG AAG AAA	861
	Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys	
	280	285
15	TTT TTG AGC ACC TTA ACT ATA GAT GGT GTA ACC CGG	897
	Phe Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg	
	290	295
20	AGT GAC CAA GGA TTG TAC ACC TGT GCA GCA TCC AGT	933
	Ser Asp Gln Gly Leu Tyr Thr Cys Ala Ala Ser Ser	
	300	305
	310	
	315	320
25	GGG CTG ATG ACC AAG AAG AAC AGC ACA TTT GTC AGG	969
	Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val Arg	
	325	330
	335	
30	GTC CAT GAA AAA CCT TTT GTT GCT TTT GGA AGT GGC	1005
	Val His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly	
	340	345
	350	355
	ATG GAA TCT CTG GTG GAA GCC ACG GTG GGG GAG CGT	1041
	Met Glu Ser Leu Val Glu Ala Thr Val Gly Glu Arg	
	355	
	355	
35	GTC AGA ATC CCT GCG AAG TAC CTT GGT TAC CCA CCC	1077
	Val Arg Ile Pro Ala Lys Tyr Leu Gly Tyr Pro Pro	

- 45 -

	CCA GAA ATA AAA TGG TAT AAA AAT GGA ATA CCC CTT	1113	
	Pro Glu Ile Lys Trp Tyr Lys Asn Gly Ile Pro Leu		
360	365	370	
5	GAG TCC AAT CAC ACA ATT AAA GCG GGG CAT GTA CTG	1149	
	Glu Ser Asn His Thr Ile Lys Ala Gly His Val Leu		
	375	380	
10	ACG ATT ATG GAA GTG AGT GAA AGA GAC ACA GGA AAT	1185	
	Thr Ile Met Glu Val Ser Glu Arg Asp Thr Gly Asn		
	385	390	395
15	TAC ACT GTC ATC CTT ACC AAT CCC ATT TCA AAG GAG	1221	
	Tyr Thr Val Ile Leu Thr Asn Pro Ile Ser Lys Glu		
	400	405	
20	AAG CAG AGC CAT GTG GTC TCT CTG GTT GTG TAT GTC	1257	
	Lys Gln Ser His Val Val Ser Leu Val Val Tyr Val		
	410	415	
25	CCA CCC CAG ATT GGT GAG AAA TCT CTA ATC TCT CCT	1293	
	Pro Pro Gln Ile Gly Glu Lys Ser Leu Ile Ser Pro		
	420	425	430
	GTG GAT TCC TAC CAG TAC GGC ACC ACT CAA ACG CTG	1329	
	Val Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu		
	435	440	
30	ACA TGT ACG GTC TAT GCC ATT CCT CCC CCG CAT CAC	1365	
	Thr Cys Thr Val Tyr Ala Ile Pro Pro Pro His His		
	445	450	455
35	ATC CAC TGG TAT TGG CAG TTG GAG GAA GAG TGC GCC	1401	
	Ile His Trp Tyr Trp Gln Leu Glu Glu Glu Cys Ala		
	460	465	

- 46 -

	AAC GAG CCC AGC CAA GCT GTC TCA GTG ACA AAC CCA	1437	
	Asn Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro		
	470	475	
5	TAC CCT TGT GAA GAA TGG AGA AGT GTG GAG GAC TTC	1473	
	Tyr Pro Cys Glu Glu Trp Arg Ser Val Glu Asp Phe		
	480	485	490
10	CAG GGA GGA AAT AAA ATT GAA GTT AAT AAA AAT CAA	1509	
	Gln Gly Gly Asn Lys Ile Glu Val Asn Lys Asn Gln		
	495	500	
15	TTT GCT CTA ATT GAA GGA AAA AAC AAA ACT GTA AGT	1545	
	Phe Ala Leu Ile Glu Gly Lys Asn Lys Thr Val Ser		
	505	510	515
20	ACC CTT GTT ATC CAA GCG GCA AAT GTG TCA GCT TTG	1581	
	Thr Leu Val Ile Gln Ala Ala Asn Val Ser Ala Leu		
	520	525	
	TAC AAA TGT GAA GCG GTC AAC AAA GTC GGG AGA GGA	1617	
	Tyr Lys Cys Glu Ala Val Asn Lys Val Gly Arg Gly		
	530	535	
25	GAG AGG GTG ATC TCC TTC CAC GTG ACC AGG GGT CCT	1653	
	Glu Arg Val Ile Ser Phe His Val Thr Arg Gly Pro		
	540	545	550
30	GAA ATT ACT TTG CAA CCT GAC ATG CAG CCC ACT GAG	1689	
	Glu Ile Thr Leu Gln Pro Asp Met Gln Pro Thr Glu		
	555	560	
35	CAG GAG AGC GTG TCT TTG TGG TGC ACT GCA GAC AGA	1725	
	Gln Glu Ser Val Ser Leu Trp Cys Thr Ala Asp Arg		
	565	570	575

- 47 -

TCT ACG TTT GAG AAC CTC ACA TGG TAC AAG CTT GGC 1761  
Ser Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly  
580 585

5 CCA CAG CCT CTG CCA ATC CAT GTG GGA GAG TTG CCC 1797  
Pro Gln Pro Leu Pro Ile His Val Gly Glu Leu Pro  
590 595

ACA CCT GTT TGC AAG AAC TTG GAT ACT CTT TGG AAA 1833  
10 Thr Pro Val Cys Lys Asn Leu Asp Thr Leu Trp Lys  
600 605 610

TTG AAT GCC ACC ATG TTC TCT AAT AGC ACA AAT GAC 1869  
Leu Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp  
15 615 620

ATT TTG ATC ATG GAG CTT AAG AAT GCA TCC TTG CAG 1905  
Ile Leu Ile Met Glu Leu Lys Asn Ala Ser Leu Gln  
625 630 635

20 GAC CAA GGA GAC TAT GTC TGC CTT GCT CAA GAC AGG 1941  
Asp Gln Gly Asp Tyr Val Cys Leu Ala Gln Asp Arg  
640 645

25 AAG ACC AAG AAA AGA CAT TGC GTG GTC AGG CAG CTC 1977  
Lys Thr Lys Lys Arg His Cys Val Val Arg Gln Leu  
650 655

30 ACA GTC CTA GAG CGT GTG GCA CCC ACG ATC ACA GGA 2013  
Thr Val Leu Glu Arg Val Ala Pro Thr Ile Thr Gly  
660 665 670

35 AAC CTG GAG AAT CAG ACG ACA AGT ATT GGG GAA AGC 2049  
Asn Leu Glu Asn Gln Thr Thr Ser Ile Gly Glu Ser  
675 680

- 48 -

	ATC GAA GTC TCA TGC ACG GCA TCT GGG AAT CCC CCT	2085	
	Ile Glu Val Ser Cys Thr Ala Ser Gly Asn Pro Pro		
	685	690	695
5	CCA CAG ATC ATG TGG TTT AAA GAT AAT GAG ACC CTT	2121	
	Pro Gln Ile Met Trp Phe Lys Asp Asn Glu Thr Leu		
	700	705	
10	GTA GAA GAC TCA GGC ATT GTA TTG AAG GAT GGG AAC	2157	
	Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn		
	710	715	
15	CGG AAC CTC ACT ATC CGC AGA GTG AGG AAG GAG GAC	2193	
	Arg Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp		
	720	725	730
20	GAA GGC CTC TAC ACC TGC CAG GCA TGC AGT GTT CTT	2229	
	Glu Gly Leu Tyr Thr Cys Gln Ala Cys Ser Val Leu		
	735	740	
	GGC TGT GCA AAA GTG GAG GCA TTT TTC ATA ATA GAA	2265	
	Gly Cys Ala Lys Val Glu Ala Phe Phe Ile Ile Glu		
	745	750	755
25	GGT GCC CAG GAA AAG ACG AAC TTG GAA ATC ATT ATT	2301	
	Gly Ala Gln Glu Lys Thr Asn Leu Glu Ile Ile Ile		
	760	765	
30	CTA GTA GGC ACG ACG GTG ATT GCC ATG TTC TTC TGG	2337	
	Leu Val Gly Thr Thr Val Ile Ala Met Phe Phe Trp		
	770	775	
35	CTA CTT CTT GTC ATC ATC CTA GGG ACC GTT AAG CGG	2373	
	Leu Leu Leu Val Ile Ile Leu Gly Thr Val Lys Arg		
	780	785	790

- 49 -

	GCC AAT GGA GGG GAA CTG AAG ACA GGC TAC TTG TCC	2409
	Ala Asn Gly Gly Glu Leu Lys Thr Gly Tyr Leu Ser	
	795	800
5	ATC GTC ATG GAT CCA GAT GAA CTC CCA TTG GAT GAA	2445
	Ile Val Met Asp Pro Asp Glu Leu Pro Leu Asp Glu	
	805	810
	815	
10	CAT TGT GAA CGA CTG CCT TAT GAT GCC AGC AAA TGG	2481
	His Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp	
	820	825
15	GAA TTC CCC AGA GAC CGG CTG AAC CTA GGT AAG CCT	2517
	Glu Phe Pro Arg Asp Arg Leu Asn Leu Gly Lys Pro	
	830	835
20	CTT GGC CGT GGT GCC TTT GGC CAA GAG ATT GAA GCA	2553
	Leu Gly Arg Gly Ala Phe Gly Gln Glu Ile Glu Ala	
	840	845
	850	
	855	860
25	GAT GCC TTT GGA ATT GAC AAG ACA GCA ACT TGC AGG	2589
	Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg	
	865	870
	875	
30	ACA GTA GCA GTC AAA ATG TTG AAA GAA GGA GCA ACA	2625
	Thr Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr	
	880	885
	890	895
35	CAC AGT GAG CAT CGA GCT CTC ATG TCT GAA CTC AAG	2661
	His Ser Glu His Arg Ala Leu Met Ser Glu Leu Lys	
	895	
	ATC CTC ATT CAT ATT GGT CAC CAT CTC AAT GTG GTC	2697
	Ile Leu Ile His Ile Gly His His Leu Asn Val Val	

- 50 -

	AAC CTT CTA GGT GCC TGT ACC AAG CCA GGA GGG CCA	2733	
	Asn Leu Leu Gly Ala Cys Thr Lys Pro Gly Gly Pro		
	900	905	910
5	CTC ATG GTG ATT GTG GAA TTC TGC AAA TTT GGA AAC	2769	
	Leu Met Val Ile Val Glu Phe Cys Lys Phe Gly Asn		
	915	920	
10	CTG TCC ACT TAC CTG AGG AGC AAG AGA AAT GAA TTT	2805	
	Leu Ser Thr Tyr Leu Arg Ser Lys Arg Asn Glu Phe		
	925	930	935
15	GTC CCC TAC AAG ACC AAA GGG GCA CGA TTC CGT CAA	2841	
	Val Pro Tyr Lys Thr Lys Gly Ala Arg Phe Arg Gln		
	940	945	
20	GGG AAA GAC TAC GTT GGA GCA ATC CCT GTG GAT CTG	2877	
	Gly Lys Asp Tyr Val Gly Ala Ile Pro Val Asp Leu		
	950	955	
	AAA CGG CGC TTG GAC AGC ATC ACC AGT AGC CAG AGC	2913	
	Lys Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser		
	960	965	970
25	TCA GCC AGC TCT GGA TTT GTG GAG GAG AAG TCC CTC	2949	
	Ser Ala Ser Ser Gly Phe Val Glu Glu Lys Ser Leu		
	975	980	
30	AGT GAT GTA GAA GAA GAG GAA GCT CCT GAA GAT CTG	2985	
	Ser Asp Val Glu Glu Glu Ala Pro Glu Asp Leu		
	985	990	995
35	TAT AAG GAC TTC CTG ACC TTG GAG CAT CTC ATC TGT	3021	
	Tyr Lys Asp Phe Leu Thr Leu Glu His Leu Ile Cys		
	1000	1005	

- 51 -

TAC AGC TTC CAA GTG GCT AAG GGC ATG GAG TTC TTG 3057  
Tyr Ser Phe Gln Val Ala Lys Gly Met Glu Phe Leu  
1010 1015

5 GCA TCG CGA AAG TGT ATC CAC AGG GAC CTG GCG GCA 3093  
Ala Ser Arg Lys Cys Ile His Arg Asp Leu Ala Ala  
1020 1025 1030

10 CGA AAT ATC CTC TTA TCG GAG AAG AAC GTG GTT AAA 3129  
Arg Asn Ile Leu Leu Ser Glu Lys Asn Val Val Lys  
1035 1040

15 ATC TGT GAC TTT GGC TTG GCC CGG GAT ATT TAT AAA 3165  
Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys  
1045 1050 1055

20 GAT CCA GAT TAT GTC AGA AAA GGA GAT GCT CGC CTC 3201  
Asp Pro Asp Tyr Val Arg Lys Gly Asp Ala Arg Leu  
1060 1065

25 CCT TTG AAA TGG ATG GCC CCA GAA ACA ATT TTT GAC 3237  
Pro Leu Lys Trp Met Ala Pro Glu Thr Ile Phe Asp  
1070 1075

30 AGA GTG TAC ACA ATC CAG AGT GAC GTC TGG TCT TTT 3273  
Arg Val Tyr Thr Ile Gln Ser Asp Val Trp Ser Phe  
1080 1085 1090

35 GGT GTT TTG CTG TGG GAA ATA TTT TCC TTA GGT GCT 3309  
Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Ala  
1095 1100

TCT CCA TAT CCT GGG GTA AAG ATT GAT GAA GAA TTT 3345  
Ser Pro Tyr Pro Gly Val Lys Ile Asp Glu Glu Phe  
1105 1110 1115

- 52 -

TGT AGG CGA TTG AAA GAA GGA ACT AGA ATG AGG GCC 3381  
Cys Arg Arg Leu Lys Glu Gly Thr Arg Met Arg Ala  
1120 1125

5 CCT GAT TAT ACT ACA CCA GAA ATG TAC CAG ACC ATG 3417  
Pro Asp Tyr Thr Thr Pro Glu Met Tyr Gln Thr Met  
1130 1135

10 CTG GAC TGC TGG CAC GGG GAG CCC AGT CAG AGA CCC 3453  
Leu Asp Cys Trp His Gly Glu Pro Ser Gln Arg Pro  
1140 1145 1150

15 ACG TTT TCA GAG TTG GTG GAA CAT TTG GGA AAT CTC 3489  
Thr Phe Ser Glu Leu Val Glu His Leu Gly Asn Leu  
1155 1160

20 TTG CAA GCT AAT GCT CAG CAG GAT GGC AAA GAC TAC 3525  
Leu Gln Ala Asn Ala Gln Gln Asp Gly Lys Asp Tyr  
1165 1170 1175

25 ATT GTT CTT CCG ATA TCA GAG ACT TTG AGC ATG GAA 3561  
Ile Val Leu Pro Ile Ser Glu Thr Leu Ser Met Glu  
1180 1185

30 GAG GAT TCT GGA CTC TCT CTG CCT ACC TCA CCT GTT 3597  
Glu Asp Ser Gly Leu Ser Leu Pro Thr Ser Pro Val  
1190 1195

35 TCC TGT ATG GAG GAG GAA GTA TGT GAC CCC AAA 3633  
Ser Cys Met Glu Glu Glu Val Cys Asp Pro Lys  
1200 1205 1210

35 TTC CAT TAT GAC AAC ACA GCA GGA ATC AGT CAG TAT 3669  
Phe His Tyr Asp Asn Thr Ala Gly Ile Ser Gln Tyr  
1215 1220

- 53 -

	CTG CAG AAC AGT AAG CGA AAG AGC CGG CCT GTG AGT	3705	
	Leu Gln Asn Ser Lys Arg Lys Ser Arg Pro Val Ser		
	1225	1230	1235
5	GTA AAA ACA TTT GAA GAT ATC CCG TTA GAA GAA CCA	3741	
	Val Lys Thr Phe Glu Asp Ile Pro Leu Glu Glu Pro		
	1240	1245	
10	GAA GTA AAA GTA ATC CCA GAT GAC AAC CAG ACG GAC	3777	
	Glu Val Lys Val Ile Pro Asp Asp Asn Gln Thr Asp		
	1250	1255	
15	AGT GGT ATG GTT CTT GCC TCA GAA GAG CTG AAA ACT	3813	
	Ser Gly Met Val Leu Ala Ser Glu Glu Leu Lys Thr		
	1260	1265	1270
20	TTG GAA GAC AGA ACC AAA TTA TCT CCA TCT TTT GGT	3849	
	Leu Glu Asp Arg Thr Lys Leu Ser Pro Ser Phe Gly		
	1275	1280	
25	GGA ATG GTG CCC AGC AAA AGC AGG GAG TCT GTG GCA	3885	
	Gly Met Val Pro Ser Lys Ser Arg Glu Ser Val Ala		
	1285	1290	1295
30	TCT GAA GGC TCA AAC CAG ACA AGC GGC TAC CAG TCC	3921	
	Ser Glu Gly Ser Asn Gln Thr Ser Gly Tyr Gln Ser		
	1300	1305	
35	GGA TAT CAC TCC GAT GAC ACA GAC ACC ACC GTG TAC	3957	
	Gly Tyr His Ser Asp Asp Thr Asp Thr Thr Val Tyr		
	1310	1315	
	TCC AGT GAG GAA GCA GAA CTT TTA AAG CTG ATA GAG	3993	
	Ser Ser Glu Glu Ala Glu Leu Leu Lys Leu Ile Glu		
	1320	1325	1330

- 54 -

ATT GGA GTG CAA ACC GGT AGC ACA GCC CAG ATT CTC 4029  
Ile Gly Val Gln Thr Gly Ser Thr Ala Gln Ile Leu  
1335 1340

5 CAG CCT GAC ACG GGG ACC ACA CTG AGC TCT CCT CCT 4065  
Gln Pro Asp Thr Gly Thr Thr Leu Ser Ser Pro Pro  
1345 1350 1355

10 GTT TAAAAGGAAG CATCCACACC CCAACTCCCCG GACATCACAT 4108  
Val  
1356

15 GAGAGGTCTG CTCAGATTGT GAAGTGTGT TCTTTCCACC 4148

AGCAGGAAGT AGCCGCATTG ATTTTCATT TCGACAACAG 4188

20 AAAAAGGACC TCGGACTGCA GGGAGCCAGC TCTTCTAGGC 4228  
TTGTGACC 4236

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 433 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

35 (ix) FEATURE:

- 55 -

(A) NAME/KEY: ckit proto-oncogene receptor

(B) LOCATION: Amino acids 543-975

5

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Yarden, Y., et al.

10 (B) JOURNAL: EMBO J.

(C) VOLUME: 6

(D) PAGES: 3341-3351

15

(E) DATE: 1987

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

20 Leu Thr Tyr Lys Tyr Leu Gln Lys Pro Met Tyr Glu Val Gln  
543 545 550 555

Trp Lys Val Val Glu Glu Ile Asn Gly Asn Asn Tyr Val Tyr  
560 565 570

25 Ile Asp Pro Thr Gln Leu Pro Tyr Asp His Lys Trp Glu Phe  
575 580

Pro Arg Asn Arg Leu Ser Phe Gly Lys Thr Leu Gly Ala Gly  
30 585 590 595

Ala Phe Gly Lys Val Val Ala Glu Thr Ala Tyr Gly Leu Ile  
600 605 610

35 Lys Ser Asp Ala Ala Met Thr Val Ala Val Lys Met Leu Lys

- 56 -

	615	620	625
	Pro Ser Ala His Leu Thr Glu Arg Glu Ala Leu Met Ser Glu		
	630	635	640
5	Leu Lys Val Leu Ser Tyr Leu Gly Asn His Met Asn Ile Val		
	645	650	
	Asn Leu Leu Gly Ala Cys Thr Ile Gly Gly Pro Thr Leu Val		
10	655	660	665
	Ile Thr Glu Tyr Cys Cys Tyr Gly Asp Leu Leu Asn Phe Leu		
	670	675	680
15	Arg Arg Lys Arg Asp Ser Phe Ile Cys Ser Lys Gln Glu Asp		
	685	690	695
	His Ala Glu Ala Ala Leu Tyr Lys Asn Leu Leu His Ser Lys		
	700	705	710
20	Glu Ser Ser Cys Ser Asp Ser Thr Asn Glu Tyr Met Asp Met		
	715	720	
	Lys Pro Gly Val Ser Tyr Val Val Pro Thr Lys Ala Asp Lys		
25	725	730	735
	Arg Arg Ser Val Arg Ile Gly Ser Tyr Ile Glu Arg Asp Val		
	740	745	750
30	Thr Pro Ala Ile Met Glu Asp Asp Glu Leu Ala Leu Asp Leu		
	755	760	765
	Glu Asp Leu Leu Ser Phe Ser Tyr Gln Val Lys Gly Met Ala		
	770	775	780

- 57 -

Phe Leu Ala Ser Lys Asn Cys Ile His Arg Asp Leu Ala Ala  
785 790

Arg Asn Ile Leu Leu Thr His Gly Arg Ile Thr Lys Ile Cys  
5 795 800 805

Asp Phe Gly Leu Ala Arg Asp Ile Lys Asn Asp Ser Asn Tyr  
810 815 820

10 Val Val Lys Gly Asn Ala Arg Leu Pro Val Lys Val Met Ala  
825 830 835

Pro Glu Ser Ile Phe Asn Cys Val Tyr Thr Glu Glu Ser Asp  
840 845 850

15 Val Trp Ser Tyr Gly Ile Phe Leu Trp Glu Leu Phe Ser Leu  
855 860

Gly Ser Ser Pro Tyr Pro Gly Met Pro Val Lys Ser Lys Phe  
20 865 870 875

Tyr Lys Met Ile Lys Glu Gly Phe Arg Met Leu Ser Pro Glu  
880 885 890

25 His Ala Pro Ala Glu Met Tyr Asp Ile Met Lys Thr Cys Trp  
895 900 905

Asp Ala Asp Pro Leu Lys Arg Pro Thr Phe Lys Gln Ile Val  
910 915 920

30 Gln Leu Ile Glu Lys Gln Ile Ser Glu Ser Thr Asn His Ile  
925 930

Tyr Ser Asn Leu Ala Asn Cys Ser Pro Asn Arg Gln Lys Pro  
35 935 940 945

- 58 -

Val Val Asp His Ser Val Arg Ile Asn Ser Val Gly Ser Thr  
950 955 960

5 Ala Ser Ser Ser Gln Pro Leu Leu Val His Asp Asp Val  
965 970 975

(2) INFORMATION FOR SEQ ID NO: 9:

10 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 437 amino acids

(B) TYPE: amino acid

15 (C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

20 (ix) FEATURE:

(A) NAME/KEY: CSF-1 receptor

25 (B) LOCATION: Amino acids 536-972

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Coussens, L., et al.

30 (B) JOURNAL: Nature

(C) VOLUME: 320

35 (D) PAGES: 277-280

- 59 -

(E) DATE: 1986

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Leu Leu Tyr Lys Tyr Lys Gln Lys Pro Lys Tyr Gln Val Arg  
536 540 545

10 Trp Lys Ile Ile Glu Ser Tyr Glu Gly Asn Ser Tyr Thr Phe  
550 555 560

Ile Asp Pro Thr Gln Leu Pro Tyr Asn Glu Lys Trp Glu Phe  
565 570 575

15 Pro Arg Asn Asn Leu Gln Phe Gly Lys Thr Leu Gly Ala Gly  
580 585 590

Ala Phe Gly Lys Val Val Glu Ala Thr Ala Phe Gly Leu Gly  
20 595 600 605

Lys Glu Asp Ala Val Leu Lys Val Ala Val Lys Met Leu Lys  
610 615

25 Ser Thr Ala His Ala Asp Glu Lys Glu Ala Leu Met Ser Glu  
620 625 630

Leu Lys Ile Met Ser His Leu Gly Gln His Glu Asn Ile Val  
635 640 645

30 Asn Leu Leu Gly Ala Cys Thr His Gly Gly Pro Val Leu Val  
650 655 660

Ile Thr Glu Tyr Cys Cys Tyr Gly Asp Leu Leu Asn Phe Leu  
35 665 670 675

- 60 -

Arg Arg Lys Ala Glu Ala Met Leu Gly Pro Ser Leu Ser Pro  
680 685

Gly Gln Asp Pro Glu Gly Gly Val Asp Tyr Lys Asn Ile His  
5 690 695 700

Leu Glu Lys Lys Tyr Val Arg Arg Asp Ser Gly Phe Ser Ser  
705 710 715

10 Gln Gly Val Asp Thr Tyr Val Glu Met Arg Pro Val Ser Thr  
720 725 730

Ser Ser Asn Asp Ser Phe Ser Glu Gln Asp Leu Asp Lys Glu  
735 740 745

15 Asp Gly Arg Pro Leu Glu Leu Arg Asp Leu Leu His Phe Ser  
750 755

Ser Gln Val Ala Gln Gly Met Ala Phe Leu Ala Ser Lys Asn  
20 760 765 770

Cys Ile His Arg Asp Val Ala Ala Arg Asn Val Leu Leu Thr  
775 780 785

25 Asn Gly His Val Ala Lys Ile Gly Asp Phe Gly Leu Ala Arg  
790 795 800

Asp Ile Met Asn Asp Ser Asn Tyr Ile Val Lys Gly Asn Ala  
805 810 815

30 Arg Leu Pro Val Lys Trp Met Ala Pro Glu Ser Ile Phe Asp  
820 825

Cys Val Tyr Thr Val Gln Ser Asp Val Trp Ser Tyr Gly Ile  
35 830 835 840

- 61 -

	Leu Leu Trp Glu Ile Phe Ser Leu Gly Leu Asn Pro Tyr Pro		
	845	850	855
	Gly Ile Leu Val Asn Ser Lys Phe Tyr Lys Leu Val Lys Asp		
5	860	865	870
	Gly Tyr Gln Met Ala Gln Pro Ala Phe Ala Pro Lys Asn Ile		
	875	880	885
10	Tyr Ser Ile Met Gln Ala Cys Trp Ala Leu Glu Pro Thr His		
	890	895	
	Arg Pro Thr Phe Gln Gln Ile Cys Ser Phe Leu Gln Glu Gln		
	900	905	910
15	Ala Gln Glu Asp Arg Arg Glu Arg Asp Tyr Thr Asn Leu Pro		
	915	920	925
	Ser Ser Ser Arg Ser Gly Gly Ser Gly Ser Ser Ser Ser Glu		
20	930	935	940
	Leu Glu Glu Glu Ser Ser Ser Glu His Leu Thr Cys Cys Glu		
	945	950	955
25	Gln Gly Asp Ile Ala Gln Pro Leu Leu Gln Pro Asn Asn Tyr		
	960	965	
	Gln Phe Cys		
	970		
30	(2) INFORMATION FOR SEQ ID NO: 10:		

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 566 amino acids

- 62 -

(B) TYPE: amino acid

(C) STRANDEDNESS:

5

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

10 (ix) FEATURE:

(A) NAME/KEY: PDGF receptor

(B) LOCATION: Amino acids 522-1087

15

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Gronwald, R., et al.

20 (B) JOURNAL: Proc. Natl. Acad. Sci.

(C) VOLUME: 85

(D) PAGES: 3435-3439

25

(E) DATE: 1988

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

30 Met Leu Trp Gln Lys Lys Pro Arg Tyr Glu Ile Arg Trp Lys  
522 525 530 535

Val Ile Glu Ser Val Ser Ser Asp Gly His Glu Tyr Ile Tyr  
540 545

35

- 63 -

Val Asp Pro Val Gln Leu Pro Tyr Asp Ser Thr Trp Glu Leu  
550 555 560

Pro Arg Asp Gln Leu Val Leu Gly Arg Thr Leu Gly Ser Gly  
5 565 570 575

Ala Phe Gly Gln Val Val Glu Ala Thr Ala His Gly Leu Ser  
580 585 590

His Ser Gln Ala Thr Met Lys Val Ala Val Lys Met Leu Lys  
10 595 600 605

Ser Thr Ala Arg Ser Ser Glu Lys Gln Ser Leu Met Ser Glu  
610 615

Leu Lys Ile Met Ser His Leu Gly Pro His Leu Asn Val Val  
15 620 625 630

Asn Leu Leu Gly Ala Cys Thr Lys Gly Pro Ile Tyr Ile  
20 635 640 645

Ile Thr Glu Tyr Cys Arg Tyr Gly Asp Leu Val Asp Tyr Leu  
650 655 660

His Arg Asn Lys His Thr Phe Leu Gln Arg His Ser Asn Lys  
25 665 670 675

His Cys Pro Pro Ser Ala Glu Leu Tyr Ser Asn Ala Leu Pro  
680 685

Val Gly Phe Ser Leu Pro Ser His Leu Asn Leu Thr Gly Glu  
30 690 695 700

Ser Asp Gly Gly Tyr Met Asp Met Ser Lys Asp Glu Ser Ile  
35 705 710 715

- 64 -

Asp Tyr Val Pro Met Leu Asp Met Lys Gly Asp Ile Lys Tyr  
720 725 730

Ala Asp Ile Glu Ser Pro Ser Tyr Met Ala Pro Tyr Asp Asn  
5 735 740 745

Tyr Val Pro Ser Ala Pro Glu Arg Thr Tyr Arg Ala Thr Leu  
750 755

Ile Asn Asp Ser Pro Val Leu Ser Tyr Thr Asp Leu Val Gly  
10 760 765 770

Phe Ser Tyr Gln Val Ala Asn Gly Met Asp Phe Leu Ala Ser  
775 780 785

Lys Asn Cys Val His Arg Asp Leu Ala Ala Arg Asn Val Leu  
15 790 795 800

Ile Cys Glu Gly Lys Leu Val Lys Ile Cys Asp Phe Gly Phe  
20 805 810 815

Ala Arg Asp Ile Met Arg Asp Ser Asn Tyr Ile Ser Lys Gly  
820 825

Ser Thr Tyr Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile  
25 830 835 840

Phe Asn Ser Leu Tyr Thr Thr Leu Ser Asp Val Trp Ser Phe  
845 850 855

Gly Ile Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Thr Pro  
30 860 865 870

Tyr Pro Glu Leu Pro Met Asn Asp Gln Phe Tyr Asn Ala Ile  
35 875 880 885

- 65 -

Lys Arg Gly Tyr Arg Met Ala Gln Pro Ala His Ala Ser Asp  
890 895

5 Glu Ile Tyr Glu Ile Met Gln Lys Cys Trp Glu Glu Lys Phe  
900 905 910

Glu Thr Arg Pro Pro Phe Ser Gln Leu Val Leu Leu Leu Glu  
915 920 925

10 Arg Leu Leu Gly Glu Gly Tyr Lys Lys Lys Tyr Gln Gln Val  
930 935 940

Asp Glu Glu Phe Leu Arg Ser Asp His Pro Ala Ile Leu Arg  
945 950 955

15 Ser Gln Ala Arg Phe Pro Gly Ile His Ser Leu Arg Ser Pro  
960 965

Leu Asp Thr Ser Ser Val Leu Tyr Thr Ala Val Gln Pro Asn  
20 970 975 980

Glu Ser Asp Asn Asp Tyr Ile Ile Pro Leu Pro Asp Pro Lys  
985 990 995

25 Pro Asp Val Ala Asp Glu Gly Leu Pro Glu Gly Ser Pro Ser  
1000 1005 1010

Leu Ala Ser Ser Thr Leu Asn Glu Val Asn Thr Ser Ser Thr  
30 1015 1020 1025

Ile Ser Cys Asp Ser Pro Leu Glu Leu Gln Glu Glu Pro Gln  
1030 1035

Gln Ala Glu Pro Glu Ala Gln Leu Glu Gln Pro Gln Asp Ser  
35 1040 1045 1050

- 66 -

Gly Cys Pro Gly Pro Leu Ala Glu Ala Glu Asp Ser Phe Leu  
1055 1060 1065

5 Glu Gln Pro Gln Asp Ser Gly Cys Pro Gly Pro Leu Ala Glu  
1070 1075 1080

Ala Glu Asp Ser Phe Leu  
1085

10 (2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 16 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TCGACGCGCG ATG GAG 16

30

35

- 67 -

We claim:

1. An isolated DNA sequence encoding the Kinase insert Domain containing Receptor.
- 5 2. The DNA sequence of Claim 1 wherein said sequence is a human gene.
3. An isolated DNA sequence comprising a DNA sequence capable of hybridizing with a DNA sequence encoding the amino acid sequence of Figure 7.
- 10 4. A method for the production of a growth factor receptor which comprises transforming a host cell with the DNA sequence of Claim 3 and culturing the host cell under conditions which result in expression of the gene by an expression vector.
- 15 5. The method of Claim 4 wherein the host cell is a bacteria, virus, yeast, insect or mammalian cell line.
- 20 6. The method of Claim 5 wherein the host cell is a COS-1 cell, NIH3T3 fibroblast or CMT-3 monkey kidney cell.
7. The method of Claim 5 where the expression vector is pcDNAltkpASP expression vector.
- 25 8. A lambda gt11 phage harboring the clone BTIII081.8 (ATCC accession number 40,931) or the clone BTIII129.5 (ATCC Accession number 40,975).
9. A plasmid pBlueScript KS which contains the clone BTIV169 (ATCC accession number 75200).
- 30 10. An isolated growth factor receptor designated the Kinase insert Domain containing Receptor.
11. The receptor of Claim 10 comprising the amino acid sequence of Figure 7.
12. The receptor of Claim 10 encoded by an isolated DNA sequence comprising a DNA sequence capable

- 68 -

of hybridizing with a DNA sequence encoding the amino acid sequence of Figure 7.

5        13. A biologically active protein fragment which retains the receptor activity of the receptor of Claim 10.

10        14. An isolated DNA sequence encoding a biologically active protein fragment which retains the receptor activity of an isolated growth factor receptor designated the Kinase insert Domain containing Receptor.

15        15. An oligonucleotide primer consisting of an oligonucleotide primer having 21 bases and having a sequence depicted for Primer 1 in Figure 2.

16. An oligonucleotide primer consisting of an oligonucleotide primer having 29 bases and having a sequence depicted for Primer 2 in Figure 2.

17. The 363 base pair product having the sequence depicted in Figure 4, or a biological equivalent of said sequence.

20

25

30

35

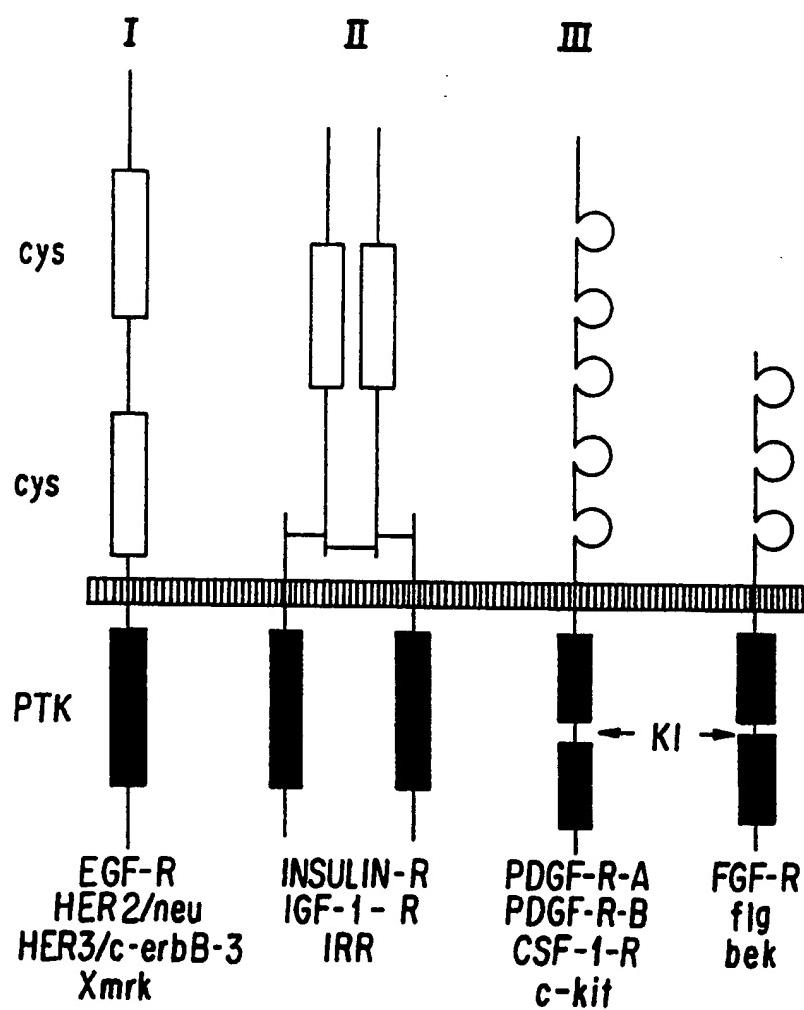


FIG. 1

2 / 28

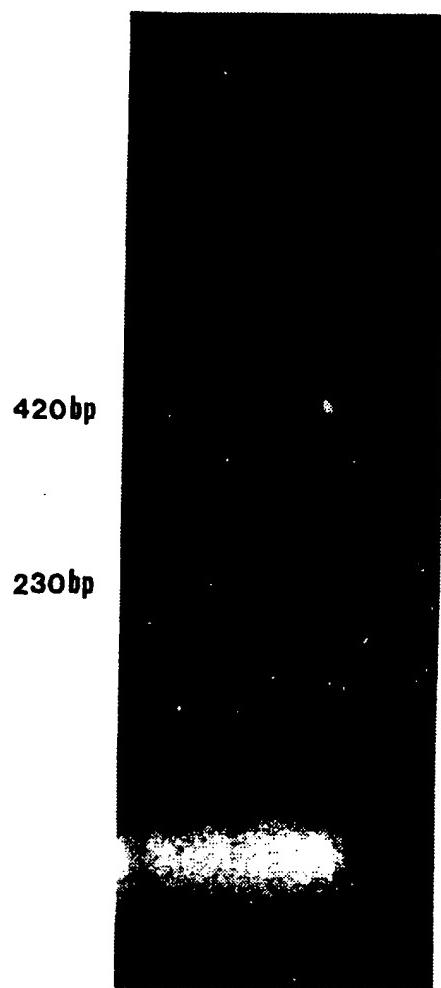
FIG. 2

<u>PRIMER 1</u>	<u>RECEPTOR</u>	<u>PRIMER 2</u>	<u>RECEPTOR</u>	<u>CONSENSUS</u>	<u>PRIMER 2</u>
PDGF ckit CSF FGF	AAC CTG TTG GGG GCC TGC ACC T A T A T C C G	GTCGAC AAC CTG TTG GGG GCC TGC AAC T A	CAC AGA GAC CTG GCG GCT AGG T G A C A C G A G C T C C C C	CAC AGA GAC CTG GCC GCT AGI C T C	GAATTCA GAG CAC GTT IGT AGC CGC CAG G T G T

WO 92/14748

PCT/US92/01300

3/28



**FIG. 3**

**SUBSTITUTE SHEET**

4/28

T C T G A G G C T C C A T G T C C A T G A G C T T G G  
A C T C A G T C G T C C T G A T G A A G G A C T T G G  
T G T C T G T C C T G C A T C A T G A A G G A C T T G G  
T G T C T G T C C T G C A T C A T G A A G G A C T T G G  
A C C T G A T G A G C A C G T C G A T G T A T C G A G A  
A C C T G A T G A G C A C G T C G A T G T A T C G A G A  
G A A G C A G G G T T G C A G T C A T G C A T G C A G A  
G A A G C A G G G T T G C A G T C A T G C A T G C A G A  
T T G A G G G T T G C A G T C A T G C A T G C A T G C A G A  
T T G A G G G T T G C A G T C A T G C A T G C A T G C A G A  
A A T G A A A C G C G C T C C A G T C A T G C A T G C A G A  
A A T G A A A C G C G C T C C A G T C A T G C A T G C A G A  
G C A A C T A C G C G C T C C A G T C A T G C A T G C A G A  
G C A A C T A C G C G C T C C A G T C A T G C A T G C A G A  
T C T G A G G A G A G C T G C A G T C T G G A G C T A G T G C T  
T C T G A G G A G A G C T G C A G T C T G G A G C T A G T G C T  
A T C A A A G A G C T G C A G T C T G G A G C T A G T G C T  
A T C A A A G A G C T G C A G T C T G G A G C T A G T G C T  
G A C T A G G A G A G C T G C A G T C T G G A G C T A G T G C T  
G A C T A G G A G A G C T G C A G T C T G G A G C T A G T G C T  
A C T G G A G A G C T G C A G T C T G G A G C T A G T G C T  
A C T G G A G A G C T G C A G T C T G G A G C T A G T G C T  
G A T G G A G A G C T G C A G T C T G G A G C T A G T G C T  
G A T G G A G A G C T G C A G T C T G G A G C T A G T G C T  
C A G C C A G G A G A G C T G C A G T C T G G A G C T A G T G C T

FIG. 4A

FIG. 4B

## SUBSTITUTE SHEET

5/28

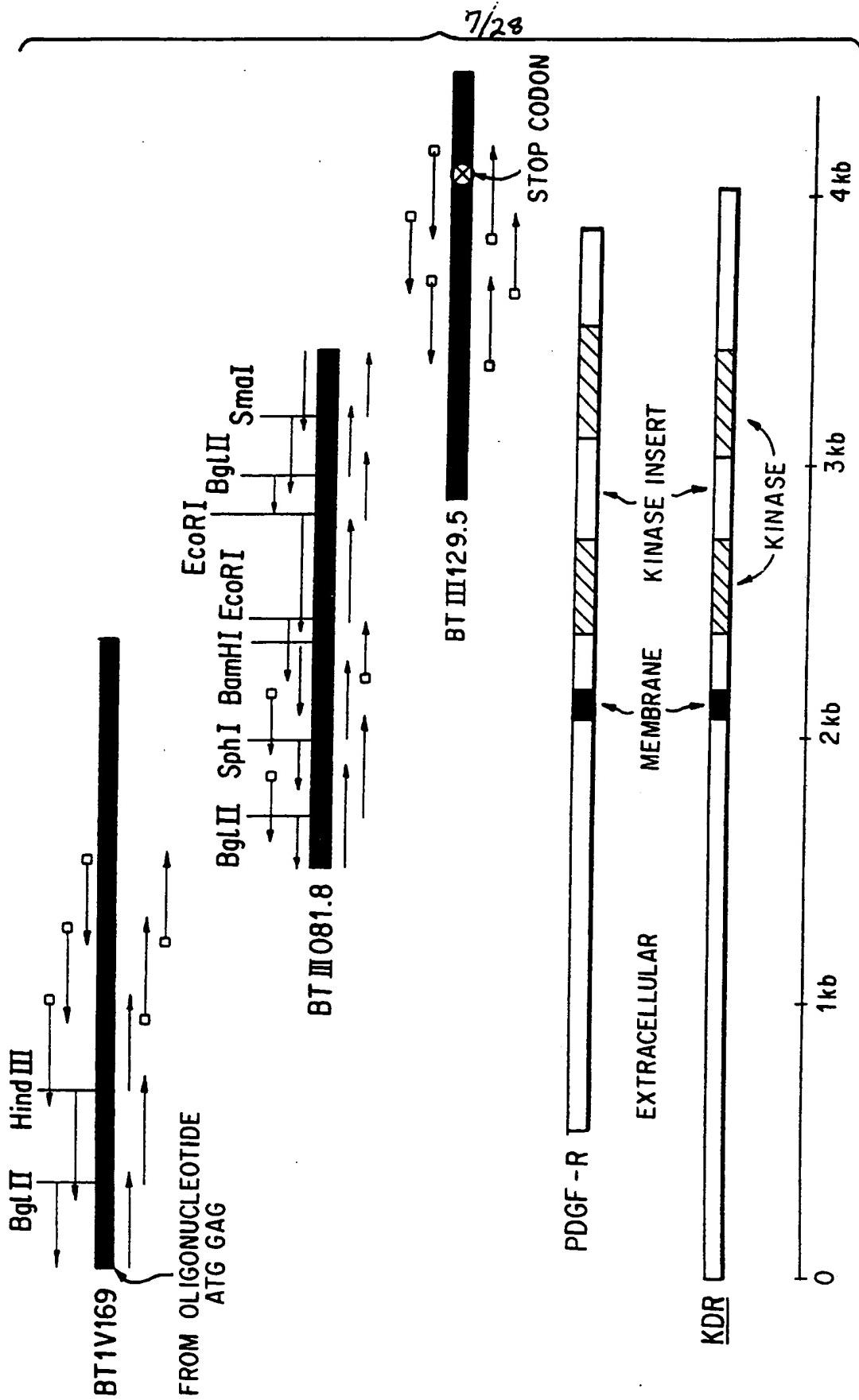
EIC.5A

6/28

FGF AAC TGC TGG GCG CCT TGAC CGAGGA TGG CCT GAT GTCA TCG GGAT ATGC CCT CAAGG CAA CCT GCG GGG GAG TACT CTG  
 AA CTG TGG GCG CCT TGAC CCT  
 230 bp AAT CTGT GGG GCG CCT TGAC AAT CGCA AAT CCT G  
 10<sup>r</sup> 20<sup>r</sup> 30<sup>r</sup> 40V 50V 60V 70V 80V  
 90V 100V 110V 120V 130V 140V 150V 160V 170V 180V  
 FGI ---AGACCCGAGGCCAGGGCTTGAAATACTGCTATAACCCAGCCACAACCAGCCAGCTCTCCCTCAAGGGACCTGGTGTCTGGCCTACCA  
 AGA G CCC C G GCT A A AGC CA GA GAGGAGC C CCA G C TGG  
 230 bp TCTAGAGTATGCCCGGGAGCTCTACAAGGAGGTGCAAGAAGAGCTGCACATTGACGAGGAGCAACGGCAGCATATGG  
 90<sup>r</sup> 100<sup>r</sup> 110<sup>r</sup> 120<sup>r</sup> 130<sup>r</sup> 140<sup>r</sup> 150<sup>r</sup>  
 190V 200V 210V 220V 230V 240V 250V  
 FGF GGA ---GGCCCGAGGCCAGGGCTTGAAAGTGCATACACGGAGACCTGGCAAGGAATGTCTG  
 GGA GGC G T GTA AAGAAG AT CAC GAGACCTGGCAAGGAAGTCTG  
 230 bp GGAGTTGGCAAGATGCTTAATGTACTGCCGTGGAAAGGTGATTCAAGAGACCTGGCAAGCAACGTGCT  
 170<sup>r</sup> 180<sup>r</sup> 190<sup>r</sup> 200<sup>r</sup> 210<sup>r</sup> 220<sup>r</sup> 230<sup>r</sup>

FIG. 5B

SUBSTITUTE SHEET



SUBSTITUTE SHEET

FIG. 6

ATG	GAG	AGC	AAG	GTG	CTG	GCC	GTC	GTC	GCC	CTG	TGG	CTC	TGC	GTG	GAG	ACC	CGG
Met	Glu	Ser	Lys	Val	Leu	Leu	Ala	Val	Ala	Leu	Trp	Leu	[Cys]	Val	Glu	Thr	Arg
60	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	50
60	60	70	70	80	80	90	90	100	100	100	100	100	100	100	100	100	50
GCC	TCC	TCT	GTG	GGT	TTG	CCT	AGT	GTT	TCT	CTT	GAT	CTG	CCC	AGG	CTC	AGC	ATA
Ala	Ala	Ser	Val	Gly	Leu	Pro	Ser	Val	Ser	Leu	Asp	Leu	Pro	Arg	Leu	Ser	Ile
110	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
110	110	120	120	130	130	140	140	150	150	150	150	150	150	150	150	150	160
CAA	AAA	GAC	ATA	CTT	ACA	ATT	AAG	GCT	AAT	ACA	ACT	CTT	CAA	ATT	ACT	TGC	AGG
Gln	Lys	Asp	Ile	Leu	Thr	Ile	Lys	Ala	Asn	Thr	Thr	Leu	Gln	Ile	Thr	[Cys]	Arg
170	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
170	170	180	180	190	190	200	200	210	210	210	210	210	210	210	210	210	210
GGA	CAG	AGG	GAC	TTG	GAC	TGG	CTT	TGG	CCC	AAT	AAT	CAG	AGT	GGC	AGT	GAG	CAA
Gly	Gln	Arg	Asp	Ile	Leu	Asp	Trp	Leu	Pro	Asn	Asn	Gln	Ser	Gly	Ser	Glu	Gln
220	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
220	220	230	230	240	240	250	250	260	260	260	260	260	260	260	260	260	270
AGG	GTG	GAG	GTG	ACT	GAG	TGC	AGC	GAT	GGC	CTC	TRC	TGT	AAG	ACA	CTC	ACA	ATT
Arg	Val	Glu	Val	Thr	Glu	[Cys]	Ser	Asp	Gly	Leu	Phe	[Cys]	Lys	Thr	Leu	Thr	Ile
280	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
280	280	290	290	300	300	310	310	320	320	320	320	320	320	320	320	320	320
CCA	AAA	GTG	ATC	GGA	AAT	GAC	ACT	GGA	GCC	TAC	AAG	TGC	TTC	TAC	CGG	GAA	ACT
Pro	Lys	Val	Ile	Gly	Asn	Asp	Thr	Gly	Ala	Tyr	Lys	[Cys]	Phe	Tyr	Arg	Glu	Thr

## **SUBSTITUTE SHEET**

FIG. 7A

330	340	350	360	370
GAC TTG GCC TCG GTC ATT TAT GTC TAT GTT CAA GAT TAC AGA TCT CCA TTT ATT Asp Leu Ala Ser Val Ile Tyr Val Tyr Val Gln Asp Tyr Arg Ser Pro Phe Ile>		*	*	*
380	390	400	410	420
GCT TCT GTT AGT GAC CAA CAT GGA GTC GTC GAG TAC ATT ACT GAG AAC AAA AAC AAA Ala Ser Val Ser Asp Gln His Gly Val Val Tyr Ile Thr Glu Asn Lys Asn Lys>	*	*	*	*
440	450	460	470	480
ACT GTG GTG ATT CCA TGT CTC GGG TCC ATT TCA AAT CTC AAC GAG TCA CTT TGT Thr val Val Ile Pro <b>Cys</b> Leu Gly Ser Ile Ser Asn Leu Asn Val Ser Leu <b>Cys</b> >	*	*	*	*
490	500	510	520	530
GCA AGA TAC CCA GAA AAG AGA ATT GTT CCT GAT GGT AAC AGA ATT TCC TGG GAC Ala Arg Tyr Pro Glu Lys Arg Phe Val Pro Asp Gly Asn Arg Ile Ser Trp Asp>	*	*	*	*
550	560	570	580	590
AGC AAG AAG GGC TTT ACT ATT CCC AGC TAC ATG ATC AGC TAT GCT GGC ATG GTC Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met Ile Ser Tyr Ala Gly Met Val>	*	*	*	*
600	610	620	630	640
TTC TGT GAA GCA AAA ATT ATT GAT GAA AGT TAC CAG TCT ATT ATG TAC ATA GTT Phe <b>Cys</b> Glu Ala Lys Ile Asn Asp Glu Ser Tyr Gln Ser Ile Met Tyr Ile Val>	*	*	*	*

SUBSTITUTE SHEET

FIG. 7B

650	660	670	680	690	700
*	*	*	*	*	*
GTC GTR GTA GGG TAT AGG ATT TAT GAT GTG GTR CTG AGT CCC TCT CAT GGA ATT					
Val Val Val Gly Tyr Arg Ile Tyr Asp Val Val Leu Ser Pro Ser His Glv Ile>					
710	720	730	740	750	
*	*	*	*	*	
GAA CTA TCT GTT GGA GAA AAG CTT GTC TTA AAT TGT ACA GCA AGA ACT GAA CTA					
Glu Leu Ser Val Gly Glu Lys Leu Val Leu Asn [Cys] Thr Ala Arg Thr Glu Leu>					
760	770	780	790	800	810
*	*	*	*	*	*
AAT GTG GGG ATT GAC RTC AAC TGG GAA TAC CCT TCT TCG AAG CAT CAG CAT AAG					
Asn Val Gly Ile Asp Phe Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys>					
820	830	840	850	860	
*	*	*	*	*	
AAA CTT GTA AAC CGA GAC CTA AAA ACC CAG TCT GGG AGT GAG ATG AAG AAA TTT & <sup>10/23</sup>					
Lys Leu Val Asn Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe>					
870	880	890	900	910	
*	*	*	*	*	
TTC AGC ACC TTA ACT ATA GAT GGT GTA ACC CGG AGT GAC CAA GGA TTG TAC ACC					
Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu Tyr Thr>					
920	930	940	950	960	
*	*	*	*	*	
TGT GCA GCA TCC AGT GGG CTG ATG ACC AAG AAC AGC ACA TTT GTC AGG GTC					
[Cys] Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val Arg Val>					

980	990	1000	1010	1020
* CAT GAA AAA CCT TTT GTR TTT GCA AGT GGC ATG GAA TCT CTG GTC GAA GCC His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly Met Glu Ser Leu Val Glu Ala>	*	*	*	*
1030	1040	1050	1060	1070
* ACG GTC GGG GAG CGT GTC AGA ATC CCT GCG AAG TAC CTT GGT TAC CCA CCC CCA Thr Val GLY Glu Arg Val Arg Ile Pro Ala Lys Tyr Leu GLY Tyr Pro Pro Pro>	*	*	*	*
1090	1100	1110	1120	1130
* GAA ATA AAA TGG TAT AAA AAT GGA ATA CCC CTT GAG TCC AAT CAC ACA ATT AAA Glu Ile Lys Trp Tyr Lys Asn Gly Ile Pro Leu Glu Ser Asn His Thr Ile Lys>	*	*	*	*
1140	1150	1160	1170	1180
* GCG GGG CAT GIA CRG ACG ATT ARG GAA GTC AGT GAA AGA GAC ACA GCA AAT TAC Ala GLY His Val Lew Thr Ile Met Glu Val Ser Glu Arg Asp Thr GLY Asn Tyr>	*	*	*	*
1190	1200	1210	1220	1230
* ACT GTC ATC CTT ACC AAT CCC ATT TCA AAG GAG CAG AGC CAT GTC GRC TCT Thr Val Ile Lew Thr Asn Pro Ile Ser Lys Glu Lys Ser His Val Val Ser>	*	*	*	*
1250	1260	1270	1280	1290
* CTG GTT GTC TAT GTC CCA CCC CAG ATT GGT GAG AAA TCT CTA ATC TCT CCT GTC Leu Val Val Tyr Val Pro Pro Gln Ile GLY Glu Lys Ser Leu Ile Ser Pro Val>	*	*	*	*

SUBSTITUTE SHEET

FIG. 7D

1300	1310	1320	*	1330	*	1340	*	1350	*
GAT TCC TAC CAG TAC GGC ACC ACT CAA ACG CTG ACA TGT ACG GTC TAT GCC ATT									
Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu Thr [Cys]									
1360	1370	1380	*	1390	*	1400	*		
CCT CCC CCG CAT CAC ATC CAC TGG TAT TGG CAG TTG GAG GAA GAG TGC GCC AAC									
Pro Pro Pro His His Ile His Trp Tyr Trp Gln Leu Glu Glu [Cys] Ala Asn>									
1410	1420	1430	*	1440	*	1450	*		
GAG CCC AGC CAA GCT GTC TCA GTG ACA AAC CCA TAC CCR TGT GAA GAA TGG AGA									
Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro Tyr Pro [Cys] Glu Glu Trp Arg>									
1460	1470	1480	*	1490	*	1500	*	1510	*
AGT GTG GAG GAC TTC CAG GGA GGA AAT AAA ATT GAA GTT AAT AAA AAT CAA TTT									
Ser Val Glu Asp Phe Gln Gly Asn Lys Ile Glu Val Asn Lys Asn Gln Phe>									
1520	1530	1540	*	1550	*	1560	*		
GCT CTA ATT GAA GGA AAA AAC AAA ACT GTA AGT ACC CTT GTR ATC CAA GCG GCA									
Ala Leu Ile Glu Gly Lys Asn Lys Thr Val Ser Thr Leu Val Ile Gln Ala Ala>									
1570	1580	1590	*	1600	*	1610	*	1620	*
* * ATT GTG TCA GCT TTG TAC AAA TGT GAA GCG GTC AAC AAA GTC GGG AGA GGA GAG									
Asn Val Ser Ala Leu Tyr Lys [Cys] Glu Ala Val Asn Lys Val Gly Arg Gly Glu>									

SUBSTITUTE SHEET

FIG. 7E

1630	1640	1650	1660	1670
*	*	*	*	*
AGG GTC ATC TCC TRC CAC GTC ACC AGG GGT CCT GAA ATT ACT TTG CAA CCT GAC				
Arg Val Ile Ser Phe His Val Thr Arg Gly Pro Glu Ile Thr Leu Gln Pro Asp>				
1680	1690	1700	1710	1720
*	*	*	*	*
ATG CAG CCC ACT GAG CAC GAG AGC GTC TCT TTG TGC ACT GCA GAC AGA TCT				
Met Gln Pro Thr Glu Gln Ser Val Ser Leu Trp [Cys] Thr Ala Asp Arg Ser>				
1730	1740	1750	1760	1770
*	*	*	*	*
ACG TTT GAG AAC CTC ACA TGG TAC AAG CTT GGC CCA CAG CCT CTG CCA ATC CAT				
Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly Pro Gln Pro Leu Pro Ile His>				
1790	1800	1810	1820	1830
*	*	*	*	*
GTC GGA GAG TTG CCC ACA CCT GTT TGC AAG AAC TTG GAT ACT CTR TGG AAA TTG				
Val Gly Glu Leu Pro Thr Pro Val [Cys] Lys Asn Leu Asp Thr Leu Trp Lys Leu>				
1840	1850	1860	1870	1880
*	*	*	*	*
AAT GCC ACC ATG TTG TCT ATT AGC ACA AAT GAC ATT TTG ATC ATG GAG CTR AAG				
Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp Ile Leu Ile Met Glu Leu Lys>				
1900	1910	1920	1930	1940
*	*	*	*	*
AAT GCA TCC TTG CAG GAC CAA GGA GAC TAT GTC TGC CTR GCT CAA GAC AGG AAG				
Asn Ala Ser Leu Gln Asp Gln Gly Asp Tyr Val Cys Leu Ala Gln Asp Arg Lys>				

1950	*	1960	*	1970	*	1980	*	1990	*		
ACC AAG AAA AGA CAT TGC GTC GTC CAG CTC ACA GTC CTA GAG CGT GTC GCA											
Thr Lys Arg His [Cys] val Val Arg Glu Leu Thr Val Leu Glu Arg Val Ala >											
2000	*	2010	*	2020	*	2030	*	2040	*	2050	*
CCC ACG ATC ACA GGA AAC CTC GAG AAT CAG ACG ACA AGT ATT GGG GAA AGC ATC											
Pro Thr Ile Thr Gly Asn Leu Glu Asn Glu Asn Gln Thr Thr Ser Ile Gly Glu Ser Ile >											
2060	*	2070	*	2080	*	2090	*	2100	*	2110	*
GAA GTC TCA TGC ACG GCA TCT GGG AAT CCC CCT CCA CAG ATC ATG TGC TTT AAA											
Glu Val Ser [Cys] Thr Ala Ser Gly Asn Pro Pro Pro Gln Ile Met Trp Phe Lys >											
2110	*	2120	*	2130	*	2140	*	2150	*	2160	*
GAT AAT GAG ACC CTT GTC GAA GAC TCA GGC ATT GTC AAG GAT GGG AAC CGG											
Asp Asn Glu Thr Leu Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn Arg >											
2170	*	2180	*	2190	*	2200	*	2210	*	2220	*
AAC CTC ACT ATC CGC AGA GTG AGG AAG GAG GAC GAA GGC CTC TAC ACC TGC CAG											
Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp Glu Gly Leu Tyr Thr [Cys] Gln >											
2230	*	2240	*	2250	*	2260	*				
GCA TGC AGT GTT CTR GGC TGT GCA AAA GTC GAG GCA TTT TRC ATA ATA GAA GGT											
Ala [Cys] Ser Val Leu Glu [Cys] Ala Lys Val Glu Ala Phe Phe Ile Ile Glu Gly >											

FIG. 7G

SUBSTITUTE SHEET

2270	2280	2290	2300	2310	2320
*	*	*	*	*	*
GCC CAG GAA AAG AAC TAC GAA		ATC ATT ATT CTA GTA GGC ACG ACG GTG ATT			
Ala Gln Glu Lys Thr Asn Leu Glu	Ile Ile Ile Ile Ile Ile	Leu Val Gly Thr Val Ile>			
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2330	2340	2350	2360	2370	
*	*	*	*	*	
GCC ATG TTC TIC TGG CTA CTT CTT GTC ATC ATC CTA GGG ACC GTT AAG CGG GCC					
Ala Met Phe Phe Trp Leu Leu Leu Val Ile Ile Ile	Leu Gly Thr Val Lys Arg Ala>				
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2380	2390	2400	2410	2420	2430
*	*	*	*	*	*
AAT GGA GGG GAA CTG AAG ACA GGC TAC TIG TCC ATC GTC ATG GAT CCA GAT GAA					
Asn Gly Gly Glu Leu Lys Thr Gly Tyr Leu Ser Ile Val Met Asp Pro Asp Glu>					
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2440	2450	2460	2470	2480	
*	*	*	*	*	
CTC CCA TTG GAT GAA CAT TGT GAA CGA CTG CCT TAT GAT GCC AGC AAA TGG GAA					
Leu Pro Leu Asp Glu His Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp Glu>					
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2490	2500	2510	2520	2530	
*	*	*	*	*	
TTC CCC AGA GAC CGG CTG AAC CTA GGT AAG CCT CCT GGC CGT GGT GCC TTT GGC					
Phe Pro Arg Asp Arg Leu Asn Leu Gly Lys Pro Leu Gly Arg Gly Ala Phe Gly>					
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
2540	2550	2560	2570	2580	2590
*	*	*	*	*	*
CAA GAG ATT GAA GCA GAT GCC TTT GGA ATT GAC AAG ACA GCA ACT TGC AGG ACA					
Gln Glu Ile Glu Ala Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg Thr>					

16/28

2600	2610	2620	2630	2640
*	*	*	*	*
GTA GCA GTC AAA ATG TTG AAA GAA GGA GCA ACA CAC AGT GAG CAT CGA GCT CTC				
Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr His Ser Glu His Arg Ala Leu>				
2650	2660	2670	2680	2690
*	*	*	*	*
ARG TCT GAA CTC AAG ATC CTC ATT CAT ATT GGT CAC CAT CTC AAT GTG GTC AAC				*
Met Ser Glu Leu Lys Ile Leu Ile His Ile Gly His His Leu Asn Val Val Asn>				
2710	2720	2730	2740	2750
*	*	*	*	*
CTT CTA GGT GCC TGT ACC AAG CCA GGA GGG CCA CTC ATG GTG ATT GTG GAA TTC				
Leu Leu Gly Ala Cys Thr Lys Pro Gly Pro Leu Met Val Ile Val Glu Phe>				
2760	2770	2780	2790	2800
*	*	*	*	*
TGC AAA TTT GGA AAC CTG TCC ACT TAC CTG AGG AGC AAG AGA AAT GAA TTT GTC				
Cys Lys Phe Gly Asn Leu Ser Thr Tyr Leu Arg Ser Lys Arg Asn Glu Phe Val>				
2810	2820	2830	2840	2850
*	*	*	*	*
CCC TAC AAG ACC AAA GGG GCA CGA TTC CGT CAA GGG AAA GAC TAC GTT GGA GCA				
Pro Tyr Lys Thr Lys Gly Ala Arg Phe Arg Gln Gly Lys Asp Tyr Val Gly Ala>				
2870	2880	2890	2900	2910
*	*	*	*	*
ATC CCT GTG GAT CTG AAA CGG CGC TTG GAC AGC ATC ACC AGT AGC CAG AGC TCA				
Ile Pro Val Asp Leu Lys Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser Ser>				

17/28

2920	2930	2940	2950	2960	2970
<u>GCC AGC TCT GGA TTT GRG GAG GAG AAG TCC CTC AGT GAT GTA GAA GAA GAG GAA</u>					
<u>Ala Ser Ser Gly Phe Val Glu Lys Ser Leu Ser Asp Val Glu Glu Glu Glu</u>					
2980	2990	3000	3010	3020	
<u>GCT CCT GAA GAT CTG TAT AAC GAC TAC CTC ACC TTG GAG CAT CTC ATC TGT TAC</u>					
<u>Ala Pro Glu Asp Leu Tyr Lys Asp Phe Leu Thr Leu Glu His Leu Ile Cys Tyr</u>					
3030	3040	3050	3060	3070	
<u>AGC TTC CAA GTG GCT AAC GGC ATG GAG TTC TTG GCA TCG CGA AAG TGT ATC CAC</u>					
<u>Ser Phe Gln Val Ala Lys Gly Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His</u>					
3080	3090	3100	3110	3120	3130
<u>AGG GAC CTG GCG GCA CGA AAT ATC CTC TTA TCG GAG AAC GTG GTG AAA ATC</u>					
<u>Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu Lys Asn Val Val Lys Ile</u>					
3140	3150	3160	3170	3180	
<u>TCT GAC TTT GCC TTG GCC CGG GAT ATT TAT AAA GAT CCA GAT TAT GTC AGA AAA</u>					
<u>Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asp Pro Asp Tyr Val Arg Lys</u>					
3190	3200	3210	3220	3230	3240
<u>GGA GAT GCT CGC CTC CCT TTG AAA TGG ATG GCC CCA GAA ACA ATT TTT GAC AGA</u>					
<u>Gly Asp Ala Arg Leu Pro Leu Pro Glu Met Ala Pro Glu Thr Ile Phe Asp Arg</u>					

SUBSTITUTE SHEET

FIG. 7J

3250	3260	3270	3280	3290
*	*	*	*	*
GTC TAC ACA ATC CAG AGT GAC GTC TGG TCT TTT GGT GTR RTG CTG TGG GAA ATA				
Val Tyr Thr Ile Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Tyr Glu Ile>				
3300	3310	3320	3330	3340
*	*	*	*	*
TTT TCC TTA GGT GCT TCT CCA TAT CCT GGG GTA AAG ATT GAT GAA GAA TTT TGT				
Phe Ser Leu Gly Ala Ser Pro Tyr Pro Gly Val Lys Ile Asp Glu Glu Phe Cys>				
3350	3360	3370	3380	3390
*	*	*	*	*
AGG CGA TTG AAA GAA GGA ACT AGA ATG AGG GCC CCT GAT TAT ACT ACA CCA GAA				
Arg Arg Leu Lys Glu Gly Thr Arg Met Arg Ala Pro Asp Tyr Thr Pro Glu>				
3410	3420	3430	3440	3450
*	*	*	*	*
ATG TAC CAG ACC ATG CTG GAC TGC TGG CAC GGG GAG CCC AGT CAG AGA CCC ACG				
Met Tyr Gln Thr Met Leu Asp Cys Trp His Glu Pro Ser Gln Arg Pro Thr>				
3460	3470	3480	3490	3500
*	*	*	*	*
TTT TCA GAG TTG GTG GAA CAT TTG GGA AAT CTC TTG CAA GCT AAT GCT CAG CAG				
Phe Ser Glu Ile Val Glu His Leu Gly Asn Ile Leu Gln Ala Asn Ala Gln Gln>				
3520	3530	3540	3550	3560
*	*	*	*	*
GAT GGC AAA GAC TAC ATT GTT CTT CCC ATA TCA GAG ACT TTG AGC ATG GAA GAG				
Asp Gly Lys Asp Tyr Ile Val Leu Pro Ile Ser Glu Thr Leu Ser Met Glu Glu>				

**SUBSTITUTE SHEET**

**FIG. 7K**

3570	*	3580	*	3590	*	3600	*	3610	*										
GAT	TCT	GGA	CRC	TCT	CTG	CCT	ACC	TCA	CCT	GTT	TCC	TGT	ATG	GAG	GAG	GAA			
Asp	Ser	Gly	Leu	Ser	Leu	Pro	Thr	Ser	Pro	Val	Ser	Cys	Met	Glu	Glu	Glu			
3620	*	3630	*	3640	*	3650	*	3660	*	3670	*								
GTA	TGT	GAC	CCC	AAA	RTC	CAT	TAT	GAC	AAC	ACA	GCA	GGA	ATC	AGT	CAG	TAT	CTG		
Val	Cys	Asp	Pro	Lys	Phe	His	Tyr	Asp	Asn	Thr	Ala	Gly	Ile	Ser	Gln	Tyr	Leu		
3680	*	3690	*	3700	*	3710	*	3720	*	3730	*	3740	*	3750	*	3760	*	3770	*
CAG	AAC	AGT	AAG	CGA	AAG	AGC	CGG	CCT	GTG	AGT	GTA	AAA	ACA	TTT	GAA	GAT	ATC		
Gln	Asn	Ser	Lys	Arg	Lys	Ser	Arg	Pro	Val	Ser	Vai	Lys	Thr	Phe	Glu	Asp	Ile		
3790	*	3800	*	3810	*	3820	*	3830	*	3840	*	3850	*	3860	*	3870	*	3880	*
GGT	ATG	GTT	CTT	GCC	TCA	GAA	GAG	CTG	AAA	ACT	TTG	GAA	GAC	AGA	ACC	AAA	TTA		
Gly	Met	Val	Leu	Ala	Ser	Glu	Glu	Leu	Lys	Thr	Leu	Glu	Asp	Arg	Thr	Lys	Leu		

3890	3900	3910	3920	3930	3940
*	*	*	*	*	*
GAA	GGC TCA AAC CAG ACA AGC GGC TAC CAG TCC GGA TAT CAC TCC GAT GAC ACA				
Glu	Gly Ser Asn Gln Thr Ser Gly Tyr Gln Ser Gly Tyr His Ser Asp Asp Thr>				
3950	3960	3970	3980	3990	4000
*	*	*	*	*	*
GAC ACC ACC GTG TAC TCC ACT GAG GAA GCA GAA CTT TTA AAG CTG ATA GAG ATT					
Asp Thr Thr Val Tyr Ser Ser Glu Glu Ala Glu Leu Leu Lys Leu Ile Glu Ile>					
4010	4020	4030	4040	4050	4060
*	*	*	*	*	*
GGA GTG CAA ACC GGT AGC ACA GCC CAG ATT CTC CAG CCT GAC ACG GGG ACC ACA					
Gly Val Gln Thr Gly Ser Thr Ala Gln Ile Leu Gln Pro Asp Thr Gly Thr Thr>					
4070					
*	*	*	*	*	
CTG AGC TCT CCT CCT GTT TAA					
Leu Ser Ser Pro Pro Val ***					

**FIG. 7M**

**SUBSTITUTE SHEET**

21/28

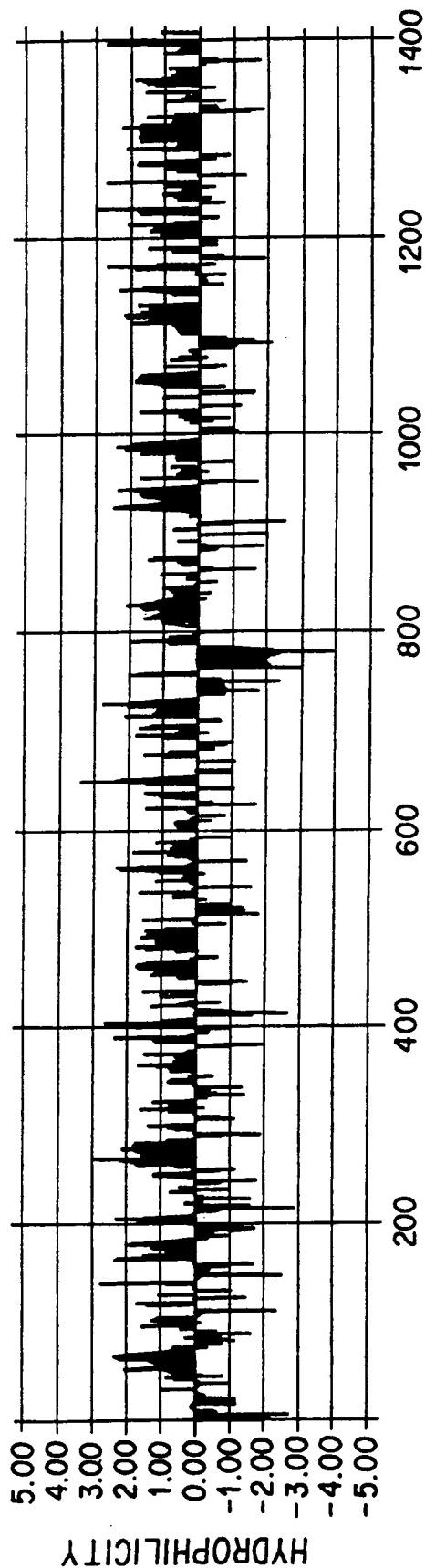


FIG. 8

SUBSTITUTE SHEET

22/28

<u>KDR</u>	787	GTVKRANGGELKTGYSIVMDPDEFLDEHCERLPYDASKWEEFPDRNLGK
<u>ckit</u>	543	L**YLQKPMYEVQWKVVEEINGNNYYIDPTQ**H-*N**SF**
CSF1	536	LLY*YKQKPKYQVRWKIIIESYEGNSYTFIDPTQ**NE-*NN*QF**
PDGF	522	MLWQKKPRYEIRWKVIESVSSDGHEYIYVDPVQ**-ST**QLV**R
 * * *		
<u>KDR</u>	839	PLGRGAFGQEIADAFGIDKTTATCRTAVAKMLKEGATHSEHRLMSELKILI
<u>ckit</u>	594	T**A**KVVAET**Y**LI**SDAAM**PS*HLT*RE**V**S
CSF1	587	T**A**KVV**T**LG**EDAVLK**ST**HAD**KE**MS
PDGF	573	T**S**VV**T**H**LSHSQATMK**ST**RSS**KQS**MS
 * * *		
<u>KDR</u>	891	HIGHHLNVVNLLGACTKPGGPLMIVEFCKFGNLSTYLRSKRNEFVPPYKTKG
<u>ckit</u>	646	YL*N*M*I**I**T**TL**T**Y**CY*D*LNF**R**DS*ICSA*QED
CSF1	639	*L*Q*E*I**I**H**V**T**Y**CY*D*LNF**R*AEMGLGPSLSP
PDGF	625	*L*P**IYI*T**Y**RY*D*VD**HRNKHT**LQRHSNK
 * * *		
<u>KDR</u>	943	ARFRQGKDYVGAIPVIDLKRRLLDSIT-SSQSSASSGFVEEKSL-----SDV
<u>ckit</u>	697	HAEA-A-L*KNLLHSKESSCS-DS*N-E-----YMDMKPGVS--YVVPT--KA
CSF1	690	GQDPPE*GVDYKN*HLEK*YVRRDSGF**GVDTYVEMRPVSTSS-NDSF*EQ
PDGF	676	HCPPSAEL*SN*LP*GFSLPSHLLNTGESDGGYMDMSKDESIDYVPMDDMKG
 * * *		
<u>KDR</u>	987	EEEAEAPDLYKDF-----LTLEHLICYSFQV
<u>ckit</u>	737	D-KRRSVRIGSYI-----ERDVTPAIMEDDELA*D*D*LSF*Y**
CSF1	741	DLDKEDGRPL-----E*RD*LHF*S**
PDGF	728	DIKY*DIESPSYMAPYDNVVP SAPERTYRATLINDSPV-*SYTD*VGF*Y**

FIG. 9A

KDR 1013 AKGMEFLASRKCIHRDLAARNILLSEKNUVKICDFGLARDIYKDPDVVRKGD

23/28

<u>KDR</u>	1065	ARLPLKWWMAPETIFDRVYTIQSDVWSFGVILWEIFSLGASPYPGVKIDEEFC
<u>ckit</u>	828	* * * V * * * S * * NC * * EE * * * Y * IF * * * S * * * MP VR SK * Y
<u>CSF1</u>	814	* * * V * * * S * * C * * V Q * * * I * * * LN * * * L V N S K * Y
<u>PDGF</u>	831	* * * S * * NS * * T L * * * I * * * GT * * * ELP MN DQ * Y
<u>ckit</u>	777	- * * A * * * N * * * * * * T H G R I T * * * * * * * * * KN * SN * * V * * N
<u>CSF1</u>	762	* Q * * A * * * N * * * * * * V * * T N G H V A * * G * * * * * * M N * SN * IV * * N
<u>PDGF</u>	779	* N * * D * * * N * * * * * * V * ICE G K L * * * * * * M R * SN * I S * * S

KDR 1117 RRLKEGTRMRAPDYTTPEMYQTMLDCWHGEPSQROPTFSELVEHLGNLLQANA  
ckit 880 KMI\*\*E\*\*LS\*EHAPA\*\*DI\*KT\*\*DAD\*LK\*\*KOIVQLEKQIEST  
CSF1 862 KLV\*D\*YQ\*AQ\*AFAPRNI\*SI\*QA\*\*AI\*\*TH\*\*QQICSF\*QEQAQEDR  
PDCF 883 NAI\*R\*Y\*\*AQ\*AHASD\*I\*EI\*QK\*\*EEKFET\*\*P\*\*Q\*\*LL\*\*ER\*\*GEGY

KDR	1169	QQDGKDYIVLPISETLSMEEDSGLSLPTSPVSCMEEEEVCDPKFHYDNTAGI
ckit	932	NHIYSNLANCSPNRQKPVVVDHSVRINSVGSTASSQPLLVHDDV
CSF1	914	RERDYTNLPPSSRSGG*GSSS*E*EEESSSEHLLTCC*QGDLIAQPLLQPNNYQ
PDGF	934	KKKYQQVDEEFLRSDHPAILR*QARF*GIHSLRSPLDTSSVLYTAVQPNESD

KDR 1213 SQYLQNSKRKSRRPVSVKRTFEDIPPLEPEVKVIPDDNQTDSGMVLASEELKTL  
CSF1 966 FC  
PDGF 987 ND<sup>a</sup> IIPPLPDPKPD \*ADEGLPPEGSPPSLASSTIENEVNTSSSTIISCDSSFL \*LQEEP

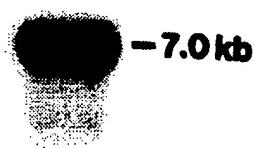
KDR 1273 EDRTKLSFPSFGGMVPSKSRESVASEGSNQTSGYQSGBYHSDDTTVYSSSEEAPDGF1.039 QQAEPEAQLEQPQDSGCPGPLAEA\*DSFLEQFPQD\*\*CPGPLAEAEADSEFL

F | G. 98 KDR 1325 ELLKLLIEIGVQTRGSTAQILLOPDGTTLSSFPV

## SUBSTITUTE SHEET

24/28

**IDENTIFICATION OF *kdp* mRNA**



**FIG. 10**

**SUBSTITUTE SHEET**

WO 92/14748

PCT/US92/01300

25/28

**IDENTIFICATION OF *kdp* GENE  
BY SOUTHERN ANALYSIS**

1 2 3 4

**FIG. 11**

SUBSTITUTE SHEET

26/28

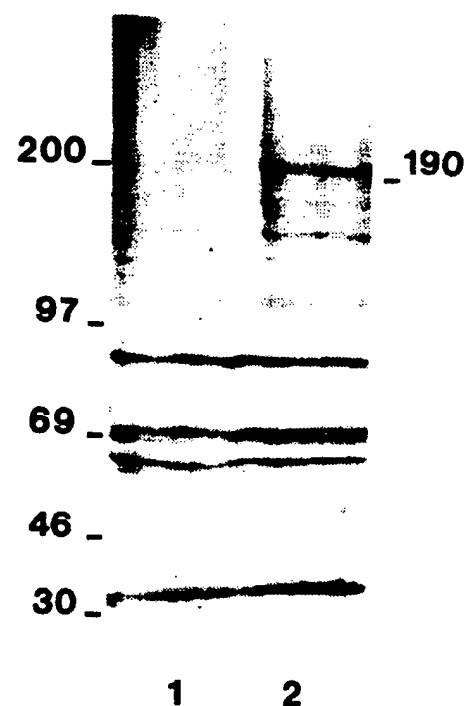


FIG. 12

27/28

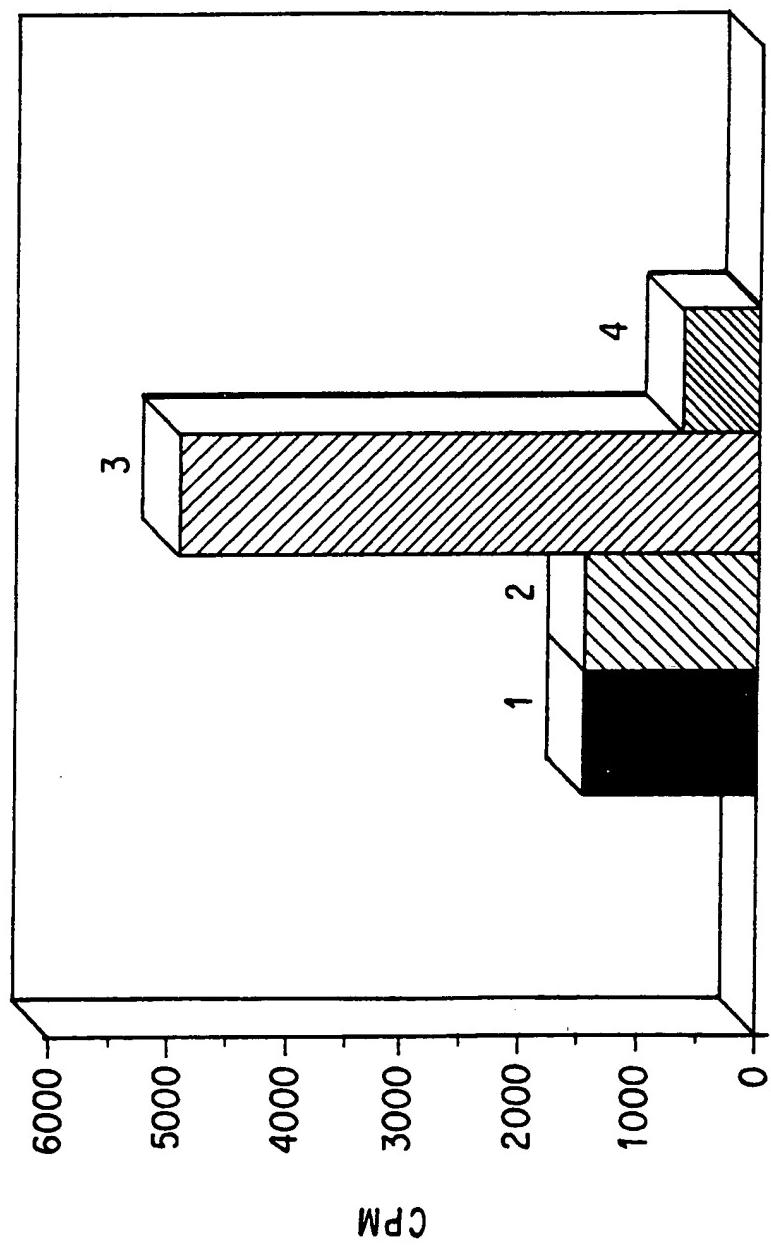


FIG. 13

SUBSTITUTE SHEET

28/28

BAND 1_	[REDACTED]	_200
BAND 2_	[REDACTED]	_97
		_69
		_46
BAND 3_	[REDACTED]	_30
		_21

1                   2

FIG. 14

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/01300

<b>I. CLASSIFICATION OF SUBJECT MATTER (If no valid classification symbols can be indicated, indicate all)<sup>3</sup></b>		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC (S): C07K 3/00, 13/00; C07H 21/00; C12P 21/06, 21/02, 21/04; C12N 15/00 US CL : 530/307; 536/27; 435/69.1, 70.1, 71.1, 320.1		
<b>II. FIELD OF SEARCHED<sup>4</sup></b>		
Minimum Documentation Searched <sup>5</sup>		
Classification System	Classification Symbols	
U.S.	530/307; 536/27; 435/69.1, 70.1, 71.1, 320.1	
Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched <sup>6</sup>		
APS, DIALOG search terms: type III receptor tyrosine kinase		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>14</sup></b>		
Category <sup>9</sup>	Character of Document <sup>10</sup> with indication, where appropriate, of the relevant portion <sup>11</sup>	Reference to Claim No. <sup>10</sup>
Y, P	Proc. Natl. Acad. Sci., Volume 88, Issued 1991, W. Mathewes et al., "A receptor tyrosine kinase cDNA isolated from a population of enriched primitive hematopoietic cells and exhibiting close genetic linkage to c-kit", pages 9026-9030, see entire document.	1-17
Y	Proc. Natl. Acad. Sci., Volume 86, Issued March 1989, A.F. Wilks, "Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction", pages 1603-1607, see entire document.	1-17
X, P	Oncogene, Volume 6, issued 1991, B.I. Terman et al., "Identification of a new endothelial cell growth factor receptor tyrosine kinase", pages 1677-1683, see entire document.	1-17
A	Oncogene, volume 3, issued 1988, M. Ruta et al., "A novel protein tyrosine kinase gene whose expression is modulated during endothelial cell differentiation", pages 9-15, see entire document.	1-17
* Special categories of cited documents: <sup>10</sup> "A" document defining the general area of the art which is not considered to be of particular relevance "B" earlier document but published on or after the International filing date "C" document which may throw doubt on priority claimed or which is cited to establish the publication date of another citation or other cited document (see notes) "D" document relating to an end disclosure, use, combination or other matter "E" document published prior to the International filing date but later than the priority date claimed "F" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "G" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step "H" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "I" document merely of interest or not directly apply		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>2</sup>	Date of Mailing of this International Search Report <sup>2</sup>	
13 MAY 1992	19 MAY 1992	
International Searching Authority <sup>1</sup>	Signature of Authorized Officer <sup>1</sup>	
ISA/US	 Lorraine M. Spector, Ph.D.	

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

A Y	Oncogene, volume 5, issued 1990, M. Shibuya et al., "Nucleotide sequence and expression of a novel human receptor-type tyrosine kinase gene (f1t) closely related to the fms family", pages 519-524, see entire document.	1-17 17
Y	Proc. Natl. Acad. Sci., Volume 85, Issued May 1988, R.G.K. Gronwald et al., "Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: Evidence for more than one receptor class", pages 3435-3439, see entire document.	15

V.  OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE<sup>1</sup>

This International Search Report has not been conducted in respect of certain claims under Article 17(2) (d) for the following reasons:

1.  Claim numbers, because they relate to subject matter (1) not required to be searched by the Authority, namely:

2.  Claim numbers, because they relate to parts of the International application that do not comply with the prescribed requirements to such an extent that no meaningful International search can be carried out (1), namely:

3.  Claim numbers, because they are dependent claims not derived in connection with the claimed and titled invention of PCT Rule 6.4(b).

VI.  OBSERVATIONS WHERE UNITY OF INVENTION IS LADING<sup>2</sup>

This International Searching Authority found multiple invasions in the International application as follows:

- I. Claims 1-9 and 14-17, drawn to nucleic acids and expression thereof. Class 536, subclass 27 and Class 435, subclass 69.1.  
II. Claims 10-13, drawn to an isolated growth factor receptor. Class 530, subclass 387.

1.  As all required additional search fees were timely paid by the applicant, the International search report covers all searchable claims of the International application. (Telepharm Franchise)

2.  As only some of the required additional search fees were timely paid by the applicant, the International search report covers only those claims of the International application for which fees were paid, specifically claims:

3.  No required additional search fees were timely paid by the applicant. Consequently, the International search report is restricted to the invention first mentioned in the claims, it is covered by claim numbers:

4.  As all searchable claims could be searched without effort justifying an additional fee, the International Search Authority did not make payment of any additional fee.

Remarks on payment:

- The additional search fees were accounted by applicant's protest.  
 No protest accounted the payment of additional search fees.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category <sup>16</sup>	Citation of Document <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
Y	Proc. Natl. Acad. Sci., Volume 86, Issued November 1989, M. Streuli et al., "A family of receptor-linked protein tyrosine phosphatases in humans and Drosophila", pages 8698-8702, see entire document.	1-14
Y	M.A. Innes et al., PCR Protocols, a guide to methods and applications, published 1990 by Academic Press (N.Y.), see page 10.	15, 16